The Annals of Applied Probability
2021, Vol. 31, No. 1, 247–283
https://doi.org/10.1214/20-AAP1589

CRAMÉR-TYPE MODERATE DEVIATION THEOREMS FOR NONNORMAL APPROXIMATION

BY QI-MAN SHAO[1,2], MENGCHEN ZHANG[3] AND ZHUO-SONG ZHANG[4,5]

[1]*Department of Statistics and Data Science, Southern University of Science and Technology*
[2]*Department of Statistics, Chinese University of Hong Kong, qmshao@sta.cuhk.edu.hk*
[3]*Department of Mathematics, Hong Kong University of Science and Technology, mzhangag@connect.ust.hk*
[4]*Department of Statistics, Chinese University of Hong Kong*
[5]*Department of Statistics and Applied Probability, National University of Singapore, zszhang.stat@gmail.com*

A Cramér-type moderate deviation theorem quantifies the relative error of the tail probability approximation. It provides a criterion whether the limiting tail probability can be used to estimate the tail probability under study. Chen, Fang and Shao (2013) obtained a general Cramér-type moderate result using Stein's method when the limiting was a normal distribution. In this paper, Cramér-type moderate deviation theorems are established for nonnormal approximation under a general Stein identity, which is satisfied via the exchangeable pair approach and Stein's coupling. In particular, a Cramér-type moderate deviation theorem is obtained for the general Curie–Weiss model and the imitative monomer-dimer mean-field model.

1. Introduction.

Consider a sequence of random variables W_n. One often needs to calculate the tail probability of W_n such as $\mathrm{P}(W_n \geq x_n)$. Since the exact distribution of W_n is hardly known, it is common to use the limiting distribution, that is, assuming that W_n converges to Y in distribution, $\mathrm{P}(Y \geq x_n)$ is used to estimate $\mathrm{P}(W_n \geq x_n)$. The Cramér-type moderate deviation seeks the largest possible a_n so that

$$\frac{\mathrm{P}(W_n \geq x)}{\mathrm{P}(Y \geq x)} = 1 + \text{error} \to 1 \tag{1.1}$$

holds for $0 \leq x \leq a_n$. This quantifies the relative error of the distribution approximation and provides a criterion whether the limiting tail probability can be used to estimate the tail probability. When Y is the normal random variable and W_n is the standardized sum of the independent random variables, the Cramér-type moderate deviation is well understood. In particular, for independent and identically distributed random variables $X_1, \dots, X_n$ with $\mathrm{E}X_i = 0$, $\mathrm{E}X_i^2 = 1$ and $\mathrm{E}e^{t_0\sqrt{|X_1|}} < \infty$, $t_0 > 0$, it holds that

$$\frac{\mathrm{P}(W_n \geq x)}{1 - \Phi(x)} = 1 + O(1)(1 + x^3)/\sqrt{n} \tag{1.2}$$

for $0 \leq x \leq n^{1/6}$, where $W_n = (X_1 + \cdots + X_n)/\sqrt{n}$. The finite-moment-generating function of $|X_1|^{1/2}$ is necessary, and both the range $0 \leq x \leq n^{1/6}$ and the order of the error term $(1 + x^3)/\sqrt{n}$ are optimal. We refer to Linnik [20] and Petrov [23], page 251, for details.

Considering general dependent random variables whose dependence is defined in terms of a Stein identity, Chen, Fang and Shao [10] obtained a general Cramér-type moderate deviation result for normal approximation using Stein's method. Stein's method, introduced by Stein [28], is a completely different approach to distribution approximation than the classical

Received September 2019; revised February 2020.
MSC2020 subject classifications. Primary 60F10; secondary 60F05.
Key words and phrases. Moderate deviation, nonnormal approximation, Stein's method, Curie–Weiss model, imitative monomer-dimer mean-field model.

Fourier transform. It works not only for independent random variables but also for dependent random variables. It can also provide accuracy of the distribution approximation. Extensive applications of Stein's method to obtain Berry–Esseen-type bounds can be found in, for example, Diaconis [16], Stein [29], Barbour [3], Goldstein and Reinert [19], Chen and Shao [12, 13], Chatterjee [6], Nourdin and Peccati [21] and Shao and Zhang [26]. We refer to Chen, Goldstein and Shao [11], Nourdin and Peccati [22] and Chatterjee [7] for comprehensive coverage of the method's fundamentals and applications. In addition to the normal approximation, Chatterjee and Shao [9] obtained a general nonnormal approximation via the exchangeable pair approach and the corresponding Berry–Esseen-type bounds. We also refer to Shao and Zhang [25] for a more general result.

The main purpose of this paper is to obtain a Cramér-type moderate deviation theorem for nonnormal approximation. Our main tool is based on Stein's method, combined with some techniques in Chatterjee and Shao [9] and Chen, Fang and Shao [10]. The paper is organized as follows. Section 2 presents a Cramér-type moderate deviation theorem under a general Stein identity setting, which recovers the result of Chen, Fang and Shao [10] as a special case. In Section 3, the result is applied to two examples: the general Curie–Weiss model and imitative monomer-dimer models. The proofs of the main results in Section 2 are given in Sections 4 and the proofs of theorems in Section 3 are postponed to Section 5.

2. Main results.

Let $W := W_n$ be the random variable of interest. Following the setting in Chatterjee and Shao [9] and Chen, Fang and Shao [10], we assume that there exists a constant δ, a nonnegative random function $\hat{K}(t)$, a function g and a random variable $R(W)$ such that

$$\mathrm{E}\big(f(W)g(W)\big) = \mathrm{E}\left(\int_{|t|\le\delta} f'(W+t)\hat{K}(t)\,dt\right) + \mathrm{E}\big(f(W)R(W)\big) \tag{2.1}$$

for all absolutely continuous functions f for which the expectation of either side exists. Let

$$\hat{K}_1 = \int_{|t|\le\delta} \hat{K}(t)\,dt \tag{2.2}$$

and

$$G(y) = \int_0^y g(t)\,dt. \tag{2.3}$$

Let Y be a random variable with the probability density function

$$p(y) = c_1 e^{-G(y)}, \quad y \in \mathbb{R}, \tag{2.4}$$

where c_1 is a normalizing constant.

In this section, we present a Cramér-type moderate deviation theorem for general distribution approximation under Stein's identity in general and under an exchangeable pair and Stein's couplings in particular.

Before presenting the main theorem, we first give some of the conditions of g.

Assume that:

(A1) The function g is nondecreasing and $g(0) = 0$.

(A2) For $y \neq 0$, $yg(y) > 0$.

(A3) There exists a positive constant c_2 such that for $x, y \in \mathbb{R}$,

$$\big|g(x+y)\big| \le c_2\big(\big|g(x)\big| + \big|g(y)\big| + 1\big). \tag{2.5}$$

(A4) There exists $c_3 \ge 1$ such that for $y \in \mathbb{R}$,

$$\big|g'(y)\big| \le c_3\left(\frac{1+|g(y)|}{1+|y|}\right). \tag{2.6}$$

A large class of functions satisfy conditions (A1)–(A4). A typical example is $g(y) = \operatorname{sgn}(y)|y|^p$, $p \ge 1$.

We are now ready to present our main theorem.

THEOREM 2.1. *Let W be a random variable of interest satisfying* (2.1). *Assume that conditions* (A1)–(A4) *are satisfied. Additionally, assume that there exist $\tau_1 > 0$, $\tau_2 > 0$, $\delta_1 > 0$ and $\delta_2 \ge 0$ such that*

$$\bigl|\mathrm{E}(\hat{K}_1 \mid W) - 1\bigr| \le \delta_1\bigl(\bigl|g(W)\bigr|^{\tau_1} + 1\bigr), \tag{2.7}$$

$$\bigl|R(W)\bigr| \le \delta_2\bigl(\bigl|g(W)\bigr|^{\tau_2} + 1\bigr). \tag{2.8}$$

In addition, there exist constants $d_0 \ge 1$, $d_1 > 0$ and $0 \le \alpha < 1$ such that

$$\mathrm{E}(\hat{K}_1 \mid W) \le d_0, \tag{2.9}$$

$$\delta\bigl|g(W)\bigr| \le d_1, \tag{2.10}$$

$$\bigl|R(W)\bigr| \le \alpha\bigl(\bigl|g(W)\bigr| + 1\bigr). \tag{2.11}$$

Then, we have

$$\begin{aligned}\frac{\mathrm{P}(W > z)}{\mathrm{P}(Y > z)} = 1 &+ O(1)\bigl(\delta\bigl(1 + zg^2(z)\bigr) \\ &+ \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr)\end{aligned} \tag{2.12}$$

for $z \ge 0$ satisfying $\delta zg^2(z) + \delta_1 zg^{\tau_1+1}(z) + \delta_2 zg^{\tau_2}(z) \le 1$, where $O(1)$ is bounded by a finite constant depending only on $d_0, d_1, c_1, c_2, c_3, \tau_1, \tau_2, \alpha$ and $\max(g(1), |g(-1)|)$.

The condition (2.1) is called a general Stein identity, see Chen, Goldstein and Shao [11], Chapter 2. We use the exchangeable pair approach and Stein's coupling to construct $\hat{K}(t)$ and $R(W)$ as follows.

Let (W, W') be an exchangeable pair, that is, (W, W') has the same joint distribution as (W', W). Let $\Delta = W - W'$. Assume that

$$\mathrm{E}(\Delta \mid W) = \lambda\bigl(g(W) - R(W)\bigr), \tag{2.13}$$

where $0 < \lambda < 1$. Assume that $|\Delta| \le \delta$ for some constant $\delta > 0$. It is known (see, e.g., Chatterjee and Shao [9]) that (2.1) is satisfied with

$$\hat{K}(t) = \frac{1}{2\lambda}\Delta\bigl(I(-\Delta \le t \le 0) - I(0 < t \le \Delta)\bigr).$$

Clearly, we have

$$\hat{K}_1 = \frac{1}{2\lambda}\Delta^2.$$

For exchangeable pairs, we have the following corollary.

COROLLARY 2.1. *For (W, W') an exchangeable pair satisfying* (2.13), *assume that $g(W)$, $\hat{K}_1$ and $R(W)$ satisfy the conditions* (A1)–(A4) *and* (2.7)–(2.11) *stated in Theorem* 2.1; *then,* (2.12) *holds.*

Stein's coupling introduced by Chen and Röllin [14] is another way to construct the general Stein identity.

A triple (W, W', T) is called a g-Stein's coupling if there is a function g such that

$$\mathrm{E}\bigl(Tf(W') - Tf(W)\bigr) = \mathrm{E}\bigl(f(W)g(W)\bigr) \tag{2.14}$$

for all absolutely continuous function f, such that the expectations on both sides exist. Assume that $|W' - W| \le \delta$. Then, by Chen and Röllin [14], we have

$$\mathrm{E}\big(f(W)g(W)\big) = \mathrm{E}\left(\int_{|t|\le\delta} f'(W+t)\hat{K}(t)\,dt\right), \tag{2.15}$$

where

$$\hat{K}(t) = T\big(I\big(0 \le t \le W' - W\big) - I\big(W' - W \le t < 0\big)\big). \tag{2.16}$$

It is easy to see that $\hat{K}_1 = T(W' - W)$.

The following corollary presents a moderate deviation result for Stein's coupling.

COROLLARY 2.2. *Let (W, W', T) be a g-Stein's coupling satisfying (2.14) and (2.15) and let $\hat{K}$ be defined as in (2.16) and assume that $\hat{K}(t) \ge 0$ for $|t| \le \delta$. Let $g(W)$ and $\hat{K}_1 := T(W' - W)$ satisfy the conditions (A1)–(A4) and (2.7), (2.9) and (2.10) stated in Theorem 2.1, then (2.12) holds with $\delta_2 = 0$.*

REMARK 2.1. For $s \ge 0$, let

$$\zeta(w,s) = \begin{cases} e^{G(w)-G(w-s)} & w > s, \\ e^{G(w)} & 0 \le w \le s, \\ 1 & w < 0. \end{cases} \tag{2.17}$$

Condition (2.7) can be replaced by

$$\big|\mathrm{E}(\hat{K}_1 \mid W) - 1\big| \le K_2 + \delta_1\big(\big|g(W)\big|^{\tau_1} + 1\big), \tag{2.18}$$

where $K_2 \ge 0$ is a random variable satisfying

$$\mathrm{E}K_2\zeta(W,s) \le \delta_1\big(1 + g^{\tau_1}(s)\big)\mathrm{E}\zeta(W,s). \tag{2.19}$$

REMARK 2.2. Condition (2.11) may not be satisfied when $|W|$ is large in some applications. Following the proof of Theorem 2.1, when (2.11) is replaced by the following condition, there exist $0 \le \alpha < 1$, $d_2 \ge 0$, $d_3 > 0$ and $\kappa > 0$ such that

$$\big|R(W)\big| \le \alpha\big(\big|g(W)\big| + 1\big) + d_2 I\big(|W| > \kappa\big), \tag{2.20}$$

and

$$d_2\mathrm{P}\big(|W| > \kappa\big) \le d_3 e^{-2s_0 d_1^{-1}\delta^{-1}}, \tag{2.21}$$

where d_1 is bounded in (2.10) and $s_0 = \max\{s : \delta s g^2(s) \le 1\}$, Theorem 2.1 and Corollaries 2.1 and 2.2 remain valid with $O(1)$ bounded by a finite constant depending only on $d_0, d_1, d_2, d_3, c_1, c_2, c_3, \tau_1, \tau_2, \alpha$ and $\max(g(1), |g(-1)|)$.

3. Applications.

In this section, we apply the main results to the general Curie–Weiss model at the critical temperature and the imitative monomer-dimer model.

3.1. *General Curie–Weiss model at the critical temperature.*

Let ξ be a random variable with probability measure ρ which is symmetric on $\mathbb{R}$. Assume that

$$\mathrm{E}\xi^2 = 1, \quad \mathrm{E}\exp\big(\beta\xi^2/2\big) < \infty \text{ for } \beta \ge 0. \tag{3.1}$$

The general Curie–Weiss model $\mathrm{CW}(\rho)$ at inverse temperature β is defined as the array of spin random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$ with joint distribution

$$d\mathrm{P}_n(\mathbf{x}) = Z_n^{-1}\exp\left(\frac{\beta}{2n}(x_1 + x_2 + \cdots + x_n)^2\right)\prod_{i=1}^n d\rho(x_i) \tag{3.2}$$

for $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ where

$$Z_n = \int \exp\left(\frac{\beta}{2n}(x_1 + x_2 + \cdots + x_n)^2\right) \prod_{i=1}^{n} d\rho(x_i)$$

is the normalizing constant.

The magnetization $m(\mathbf{x})$ is defined by

$$m(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Following the setting of Chatterjee and Dey [8], we assume that the measure ρ satisfies the following conditions:

(B1) ρ has compact support, that is, $\rho([-L, L]) = 1$ for some $L < \infty$.

(B2) Let

$$h(s) := \frac{s^2}{2} - \log \int \exp(sx)\, d\rho(x). \tag{3.3}$$

The equation $h'(s) = 0$ has a unique root at $s = 0$.

(B3) Let $k \geq 2$ be such that $h^{(i)}(0) = 0$ for $0 \leq i \leq 2k - 1$ and $h^{(2k)}(0) > 0$.

Specially, for the simple Curie–Weiss model, where $\rho = \frac{1}{2}\boldsymbol{\delta}_1 + \frac{1}{2}\boldsymbol{\delta}_{-1}$ and $\boldsymbol{\delta}$ is the Dirac measure, conditions (B1)–(B3) are satisfied with $L = 1$ and $k = 2$. For $0 < \beta < 1$, $n^{1/2}m(\mathbf{X})$ converges weakly to a Gaussian distribution, see Ellis and Newman [17]. Also, Chen, Fang and Shao [10] obtained the Cramér-type moderate deviation for this normal approximation. When $\beta = 1$, Simon and Griffiths [27] proved that the law of $n^{1/4}m(\mathbf{X})$ converges to $\mathcal{W}(4, 12)$ as $n \to \infty$, with the probability density function

$$f_Y(y) = \frac{\sqrt{2}}{3^{1/4}\Gamma(1/4)} e^{-\frac{y^4}{12}}. \tag{3.4}$$

Chatterjee and Shao [9] showed that the Berry–Esseen bound is of order $O(n^{-1/2})$.

For the rest of this subsection, we consider only the case where $\beta = 1$. Assume that conditions (B1)–(B3) are satisfied. Let $W = n^{\frac{1}{2k}} m(\mathbf{X})$. Ellis and Newman [17] showed that W converges weakly to a distribution with density

$$p(y) = c_1 \exp\bigl(-h^{(2k)}(0) y^{2k}/(2k)!\bigr), \tag{3.5}$$

where c_1 is a normalizing constant. For the concentration inequality, Chatterjee and Dey [8] used Stein's method to prove that for any $n \geq 1$ and $t \geq 0$,

$$\mathrm{P}\bigl(|W| \geq t\bigr) \leq 2e^{-c_\rho t^{2k}},$$

where $c_\rho > 0$ is a constant depending only on ρ. Moreover, Shao and Zhang [26] proved the Berry–Esseen bound:

$$\sup_{z \in \mathbb{R}} \bigl|\mathrm{P}(W \leq z) - \mathrm{P}(Y \leq z)\bigr| \leq C n^{-\frac{1}{2k}}, \tag{3.6}$$

where $Y \sim p(y)$ as defined in (3.5) and $C > 0$ is a constant.

In this subsection, we provide the Cramér-type moderate deviation for W.

THEOREM 3.1. *Let W be defined as above. If $\beta = 1$, we have*

$$\frac{\mathrm{P}(W > z)}{\mathrm{P}(Y > z)} = 1 + O(1) n^{-1/k}\bigl(1 + z^{2k+2}\bigr),$$

uniformly in $z \in (0, n^{\frac{1}{k(2k+2)}})$.

COROLLARY 3.1. *For the simple Curie–Weiss model, in which case* $\rho = \frac{1}{2}\boldsymbol{\delta}_1 + \frac{1}{2}\boldsymbol{\delta}_{-1}$ *and* $\boldsymbol{\delta}$ *is the Dirac measure. Then,*

$$\frac{\mathrm{P}(W > z)}{\mathrm{P}(Y > z)} = 1 + O(1)n^{-1/2}(1 + z^6),$$

uniformly in $z \in (0, n^{1/12})$, *where* $Y \sim \mathcal{W}(4, 12)$.

After we finished this paper, we learnt that Can and Pham [4] proved Corollary 3.1 by a completely different approach.

REMARK 3.1. Comparing to Shao and Zhang [26], Theorem 3.2(ii), we assume the additional condition that ρ is a symmetric measure. Following the proofs of Theorem 3.1 and Shao and Zhang [26], Theorem 3.2, we have (3.6) can be improved to

$$\sup_{z\in\mathbb{R}} \bigl|\mathrm{P}(W \le z) - \mathrm{P}(Y \le z)\bigr| \le Cn^{-1/k}.$$

3.2. *The imitative monomer-dimer mean-field model.* In this subsection, we consider the imitative monomer-dimer model and give the moderate deviation result. A pure monomer-dimer model can be used to study the properties of diatomic oxygen molecules deposited on tungsten or liquid mixtures with molecules of unequal size, see [18, 24] for example. Chang [5] studied the attractive component of the van der Waals potential, while Alberici, Contucci, Fedele and Mingione [1] and Alberici, Contucci and Mingione [2] considered the asymptotic properties.

Chen [15] recently obtained the Berry–Esseen bound by using Stein's method. In this subsection, we apply our main theorem to obtain the moderate deviation result.

For $n \ge 1$, let $G = (V, E)$ be a complete graph with vertex set $V = \{1, \dots, n\}$ and edge set $E = \{uv = \{u, v\} : u, v \in V, u < v\}$. A dimer configuration on the graph G is a set D of pairwise nonincident edges satisfying the following rule: if $uv \in D$, then for all $w \ne v$, $uw \notin D$. Given a dimer configuration D, the set of monomers $\mathcal{M}(D)$ is the collection of dimer-free vertices. Let $\mathbf{D}$ denote the set of all dimer configurations. Denote the number of elements by $\#(\cdot)$. Then, we have

$$2\#(D) + \#\bigl(\mathcal{M}(D)\bigr) = n.$$

We now introduce the imitative monomer-dimer model. The Hamiltonian of the model with an imitation coefficient $J \ge 0$ and an external field $h \in \mathbb{R}$ is given by

$$-T(D) = n\bigl(Jm(D)^2 + bm(D)\bigr)$$

for all $D \in \mathbf{D}$, where $m(D) = \#(\mathcal{M}(D))/n$ is called the monomer density and the parameter b is given by

$$b = \frac{\log n}{2} + h - J.$$

The associated Gibbs measure is defined as

$$p(D) = \frac{e^{-T(D)}}{\sum_{D\in\mathbf{D}} e^{-T(D)}}.$$

Let

$$H(x) = -Jx^2 - \frac{1}{2}\bigl(1 - g\bigl(\tau(x)\bigr) + \log\bigl(1 - g\bigl(\tau(x)\bigr)\bigr)\bigr), \tag{3.7}$$

where

$$g(x) = \frac{1}{2}\big(\sqrt{e^{4x} + 4e^{2x}} - e^{2x}\big), \qquad \tau(x) = (2x-1)J + h.$$

Alberici, Contucci and Mingione [2] showed that the imitative monomer-dimer model exhibits the following three phases. Let

$$J_c = \frac{1}{4(3-2\sqrt{2})}, \qquad h_c = \frac{1}{2}\log(2\sqrt{2}-2) - \frac{1}{4}.$$

There exists a function $\gamma : (J_c, \infty) \to \mathbb{R}$ with $\gamma(J_c) = h_c$ such that if $(J,h) \notin \Gamma$, where $\Gamma := \{(J, \gamma(J)) : J > J_c\}$, then the function $H(x)$ has a unique maximizer m_0 that satisfies $m_0 = g(\tau(m_0))$. Moreover, if $(J,h) \notin \Gamma \cup \{(J_c, h_c)\}$, then $H''(m_0) < 0$. If $(J,h) = (J_c, h_c)$, then $m_0 = m_c := 2 - \sqrt{2}$ and

$$H'(m_c) = H''(m_c) = H^{(3)}(m_c) = 0,$$

but

$$H^{(4)}(m_c) < 0.$$

If $(J,h) \in \Gamma$, then $H(s)$ has two distinct maximizers; therefore, in this case, $m(D)$ may not converge. Hence, we consider only the cases when $(J,h) \notin \Gamma$.

Alberici, Contucci and Mingione [2] showed that when $(J,h) \notin \Gamma \cup \{(J_c, h_c)\}$, $n^{1/2}(m(D) - m_0)$ converges to a normal distribution with zero mean and variance $\lambda_0 = -(H''(m_0))^{-1} - (2J)^{-1}$. However, when $(J,h) = (J_c, h_c)$, $n^{1/4}(m(D) - m_0)$ converges to Y in distribution, whose p.d.f. is given by

$$p(y) = c_1 e^{-\lambda_c y^4/24} \tag{3.8}$$

with $\lambda_c = -H^{(4)}(m_c) > 0$ and c_1 is a normalizing constant. Chen [15] obtained the Berry–Esseen bound using Stein's method.

We use the following notation. Let $\Sigma = \{0, 1\}^n$. For each $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma$, define a Hamiltonian

$$-T(\sigma) = n\big(Jm(\sigma)^2 + bm(\sigma)\big),$$

where $m(\sigma) = n^{-1}(\sigma_1 + \cdots + \sigma_n)$ is the magnetization of the configuration σ. Denote by $\boldsymbol{A}(\sigma)$ the set of all sites $i \in V$ such that $\sigma_i = 1$. Also, let $D(\sigma)$ denote the total number of dimer configurations $D \in \mathbf{D}$ with $\mathcal{M}(D) = \boldsymbol{A}(\sigma)$. Therefore, the Gibbs measure can be written as

$$p(\sigma) = \frac{D(\sigma)\exp(-T(\sigma))}{\sum_{\tau\in\Sigma} D(\tau)\exp(-T(\tau))}.$$

The following result gives a Cramér-type moderate deviation for the magnetization.

THEOREM 3.2. *If $(J,h) \notin \Gamma \cup \{J_c, h_c\}$, then, for $0 \le z \le n^{1/6}$,*

$$\frac{\mathrm{P}(n^{1/2}(m(\sigma) - m_0) > z)}{\mathrm{P}(Z_0 > z)} = 1 + O(1)n^{-1/2}(1+z^3), \tag{3.9}$$

where Z_0 follows normal distribution with zero mean and variance $\lambda_0 = -(H''(m_0))^{-1} - (2J)^{-1}$. If $(J,h) = (J_c, h_c)$, then for $0 \le z \le n^{1/20}$,

$$\frac{\mathrm{P}(n^{1/4}(m(\sigma) - m_c) > z)}{\mathrm{P}(Y > z)} = 1 + O(1)n^{-1/4}(1+z^5), \tag{3.10}$$

where Y is a random variable with the probability density function given in (3.8).

4. Proofs of main results. In this section, we give the proofs of the main theorems. In what follows, we use C or $C_1, C_2, \dots$ to denote a finite constant depending only on $c_1, c_2, c_3, d_0, d_1, \tau_1, \tau_2, \mu_1$ and α, where $\mu_1 = \max(g(1), |g(-1)|) + 1$, and C might be different in different places.

4.1. *Proof of Theorem* 2.1. Let Y be a random variable with a probability density function given in (2.4) and $F(z)$ be the distribution function of Y. We start with a preliminary lemma on the properties of $(1 - F(w))/p(w)$ and $F(w)/p(w)$, whose proof is postponed to Section 4.2.

LEMMA 4.1. *Assume that conditions (A1)–(A4) are satisfied. Then, we have*

$$\frac{1}{\max(1, c_3)(1 + g(w))} \le \frac{1 - F(w)}{p(w)} \le \min\left\{\frac{1}{g(w)}, 1/c_1\right\} \quad \text{for } w > 0 \tag{4.1}$$

and

$$\frac{F(w)}{p(w)} \le \min\left\{\frac{1}{|g(w)|}, 1/c_1\right\} \quad \text{for } w < 0. \tag{4.2}$$

Let f_z be the solution to Stein's equation

$$f'(w) - f(w)g(w) = I(w \le z) - F(z). \tag{4.3}$$

As shown in Chatterjee and Shao [9], the solution f_z can be written as

$$f_z(w) = \begin{cases} \dfrac{F(w)(1 - F(z))}{p(w)} & w \le z; \\ \dfrac{F(z)(1 - F(w))}{p(w)} & w > z. \end{cases} \tag{4.4}$$

Let

$$\begin{aligned} I_1 &= \mathrm{E}\left(\int_{|t| \le \delta} \left| f_z(W + t)g(W + t) - f_z(W)g(W)\right| \hat{K}(t)\,dt\right), \\ I_2 &= \mathrm{E}\left(\left|\left(\mathrm{E}(\hat{K}_1 \mid W) - 1\right) f_z(W)g(W)\right|\right), \\ I_3 &= \mathrm{E}\left(\left|\left(\mathrm{E}(\hat{K}_1 \mid W) - 1\right)\left(P(Y > z) - I(W > z + \delta)\right)\right|\right), \\ I_4 &= \mathrm{E}\left(f_z(W)\left|R(W)\right|\right). \end{aligned} \tag{4.5}$$

The following propositions provide estimates of I_1, I_2, I_3 and I_4, whose proofs are given in Section 4.4.

PROPOSITION 4.1. *If* $\delta \le 1$, *then*

$$I_1 \le C\delta. \tag{4.6}$$

Assume that $z \ge 0$, $\max(\delta, \delta_1, \delta_2) \le 1$ *and* $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1$. *Then, we have*

$$I_1 \le C\delta\left(1 + zg^2(z)\right)\left(1 - F(z)\right). \tag{4.7}$$

PROPOSITION 4.2. *We have*

$$I_2 + I_3 \le C\delta_1, \qquad I_4 \le C\delta_2. \tag{4.8}$$

For $z > 0$, $\max(\delta, \delta_1, \delta_2) \le 1$ *and* $\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1$, *we have*

$$I_2 + I_3 \le C\delta_1\left(1 + zg^{\tau_1+1}(z)\right)\left(1 - F(z)\right), \tag{4.9}$$

$$I_4 \le C\delta_2\left(1 + zg^{\tau_2}(z)\right)\left(1 - F(z)\right). \tag{4.10}$$

We are ready to give the proof of Theorem 2.1.

PROOF OF THEOREM 2.1. From (2.1), we have

$$
\begin{aligned}
\mathrm{E}&\big(f_z(W)g(W) - f_z(W)R(W)\big) \\
&= \mathrm{E}\biggl(\int_{|t|\le\delta} f_z'(W+t)\hat{K}(t)\,dt\biggr) \\
&= \mathrm{E}\biggl(\int_{|t|\le\delta} \big(f_z(W+t)g(W+t) \\
&\quad + \mathrm{P}(Y>z) - I(W+t>z)\big)\hat{K}(t)\,dt\biggr) \\
&\le \mathrm{E}\biggl(\int_{|t|\le\delta} \big(f_z(W+t)g(W+t) - f_z(W)g(W)\big)\hat{K}(t)\,dt\biggr) \\
&\quad + \mathrm{E}\big(\hat{K}_1 f_z(W)g(W)\big) \\
&\quad + \mathrm{E}\big(\hat{K}_1\big(\mathrm{P}(Y>z) - I(W>z+\delta)\big)\big) \\
&\le \mathrm{E}\biggl(\int_{|t|\le\delta} \big|f_z(W+t)g(W+t) - f_z(W)g(W)\big|\hat{K}(t)\,dt\biggr) \\
&\quad + \mathrm{E}\big(\hat{K}_1 f_z(W)g(W)\big) \\
&\quad + \mathrm{E}\big(\big|\mathrm{E}(\hat{K}_1 \mid W) - 1\big|\big|P(Y>z) - I(W>z+\delta)\big|\big) \\
&\quad + \mathrm{P}(Y>z) - \mathrm{P}(W>z+\delta).
\end{aligned} \tag{4.11}
$$

Rearranging (4.11) leads to

$$\mathrm{P}(W>z+\delta) - \mathrm{P}(Y>z) \le I_1 + I_2 + I_3 + I_4, \tag{4.12}$$

where I_1, I_2, I_3 and I_4 are defined as in (4.5).

First, we use (4.12) and Propositions 4.1 and 4.2 to prove the Berry–Esseen bound

$$\big|\mathrm{P}(W>z) - \mathrm{P}(Y>z)\big| \le C(\delta + \delta_1 + \delta_2), \tag{4.13}$$

where $C \ge 1$. By (4.12), (4.6) and (4.8), for $\delta \le 1$, we have

$$\mathrm{P}(W>z+\delta) - P(Y>z) \le C(\delta + \delta_1 + \delta_2). \tag{4.14}$$

Together with

$$\mathrm{P}(Y>z) - \mathrm{P}(Y>z+\delta) \le c_1 \int_z^{z+\delta} e^{-G(w)}\,dw \le c_1\delta,$$

we have

$$\mathrm{P}(W>z) - \mathrm{P}(Y>z) \le C(\delta + \delta_1 + \delta_2).$$

Similarly, we have

$$\mathrm{P}(W>z) - \mathrm{P}(Y>z) \ge -C(\delta + \delta_1 + \delta_2).$$

This proves the inequality (4.13) for $\delta \le 1$. For $\delta > 1$, (4.13) is trivial because $C \ge 1$.

Next, we move to prove (2.12). Let $z_0 > 1$ be a constant such that

$$\min\{z_0 g^2(z_0), z_0 g^{\tau_1+1}(z_0), z_0 g^{\tau_2}(z_0), z_0\} \ge 1.$$

For $0 \le z \le z_0$, (2.12) follows from (4.13) because

$$\frac{\mathrm{P}(W > z) - \mathrm{P}(Y > z)}{\mathrm{P}(Y > z)} \le \frac{C(\delta + \delta_1 + \delta_2)}{1 - F(z_0)}, \tag{4.15}$$

where C is a constant.

For $z > z_0$, and thus $z > 1$, we can assume $\max\{\delta, \delta_1, \delta_2\} \le 1$; otherwise, it would contradict the condition

$$\delta z g^2(z) + \delta_1 z g^{\tau_1+1}(z) + \delta_2 z g^{\tau_2}(z) \le 1. \tag{4.16}$$

In this case, it follows that

$$\delta \le 1, \delta g^2(z) \le \delta z g^2(z) \le 1, \tag{4.17}$$

provided that (4.16) holds.

By (4.12) and Propositions 4.1 and 4.2,

$$\begin{aligned} &\mathrm{P}(W > z + \delta) - \bigl(1 - F(z)\bigr) \\ &\quad \le I_1 + I_2 + I_3 + I_4 \\ &\quad \le C\bigl(1 - F(z)\bigr)\bigl(\delta\bigl(1 + zg^2(z)\bigr) + \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr). \end{aligned} \tag{4.18}$$

By replacing z with $z - \delta$, and noting that g is nondecreasing, we can rewrite (4.18) as

$$\begin{aligned} &\mathrm{P}(W > z) - \bigl(1 - F(z - \delta)\bigr) \\ &\quad \le C\bigl(1 - F(z - \delta)\bigr)\bigl(\delta\bigl(1 + zg^2(z)\bigr) + \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr). \end{aligned} \tag{4.19}$$

As $p(y)$ is decreasing in $[z - \delta, z]$, we have

$$\begin{aligned} F(z) - F(z - \delta) &= \int_{z-\delta}^{z} p(t)\, dt \\ &\le \delta p(z - \delta) \le e^{\delta g(z)} \delta p(z). \end{aligned}$$

By (4.17), it follows that $\delta g(z) \le (1/2)\delta(1 + g^2(z)) \le 1$. By (4.1), we also have

$$p(z) \le \max(1, c_3)\bigl(1 + g(z)\bigr)\bigl(1 - F(z)\bigr);$$

then,

$$F(z) - F(z - \delta) \le C\delta\bigl(1 + g(z)\bigr)\bigl(1 - F(z)\bigr)$$

for some constant C. Recall that $\delta(1 + g(z)) \le 2$; then,

$$1 - F(z - \delta) \le C\bigl(1 - F(z)\bigr).$$

Together with (4.19), we get

$$\begin{aligned} &\mathrm{P}(W > z) - \bigl(1 - F(z)\bigr) \\ &\quad \le \mathrm{P}(W > z) - \bigl(1 - F(z - \delta)\bigr) + F(z) - F(z - \delta) \\ &\quad \le C\bigl(1 - F(z - \delta)\bigr)\bigl(\delta\bigl(1 + zg^2(z)\bigr) + \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr) \\ &\qquad + C\delta\bigl(1 + g(z)\bigr)\bigl(1 - F(z)\bigr) \\ &\quad \le C\bigl(1 - F(z)\bigr)\bigl(\delta\bigl(1 + zg^2(z)\bigr) + \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr). \end{aligned}$$

Similarly, we can prove the lower bound as follows:

$$\begin{aligned} &\mathrm{P}(W > z) - \bigl(1 - F(z)\bigr) \\ &\quad \ge -C\bigl(1 - F(z)\bigr)\bigl(\delta\bigl(1 + zg^2(z)\bigr) + \delta_1\bigl(1 + zg^{\tau_1+1}(z)\bigr) + \delta_2\bigl(1 + zg^{\tau_2}(z)\bigr)\bigr). \end{aligned}$$

This completes the proof of Theorem 2.1. $\square$

4.2. *Proof of Lemma* 4.1. For $w \geq 0$, by the monotonicity of $g(\cdot)$, we have

$$\begin{aligned}1 - F(w) &= \int_w^\infty p(t)\,dt \\ &= c_1 \int_w^\infty e^{-G(t)}\,dt \\ &= c_1 \int_w^\infty \frac{1}{g(t)} e^{-G(t)}\,dG(t) \\ &\leq \frac{c_1}{g(w)} e^{-G(w)} \\ &= \frac{p(w)}{g(w)}.\end{aligned}$$

Let $H(w) = 1 - F(w) - p(w)/c_1$; then,

$$H'(w) = p(w)\bigl(g(w)/c_1 - 1\bigr).$$

Note that $g(w)/c_1 = 1$ has at most one solution in $(0, +\infty)$ and that $g(0) = 0$; then, $H(w)$ takes the maximum at either 0 or $+\infty$. We have

$$H(w) \leq \max\Bigl\{H(0), \lim_{w\to\infty} H(w)\Bigr\} \leq 0.$$

This proves the upper bound of (4.1). The inequality (4.2) can be obtained similarly.

To finish the proof, we need to prove that for $w \geq 0$,

$$\frac{p(w)}{1 + g(w)} \leq \max(1, c_3)\bigl(1 - F(w)\bigr). \tag{4.20}$$

Let

$$\zeta(w) = \frac{1}{1 + g(w)} e^{-G(w)}. \tag{4.21}$$

As $g'(w) \leq c_3(1 + g(w))$, we have

$$-\zeta'(w) = \frac{g(w)}{1 + g(w)} e^{-G(w)} + \frac{g'(w)}{(1 + g(w))^2} e^{-G(w)} \leq \max(1, c_3) e^{-G(w)}.$$

As $g(w)$ is nondecreasing and $g(w) > 0$ for $w > 0$, then $G(w) = \int_0^w g(t)\,dt \to \infty$ as $w \to \infty$. Therefore, $\lim_{w\to\infty} p(w) = 0$. Taking the integration on both sides yields

$$\zeta(w) = -\int_w^\infty \zeta'(t)\,dt \leq \max(1, c_3) \int_w^\infty e^{-G(t)}\,dt,$$

which leads to (4.20). This completes the proof.

4.3. *Preliminary lemmas*. To prove Propositions 4.1 and 4.2, we first present some preliminary lemmas. Throughout this subsection, we assume that conditions (A1)–(A4) are satisfied.

LEMMA 4.2. *Assume that* $0 < \delta \leq 1$. *Then, we have*

$$\sup_{|t|\leq\delta} \bigl|g(w + t)\bigr| \leq c_2\bigl(\bigl|g(w)\bigr| + \mu_1\bigr), \tag{4.22}$$

where $\mu_1 = \max(g(1), |g(-1)|) + 1$.

Also, for $w > s > 0$ and any positive number $a > 1$, there exists $b(a)$ depending on a, c_2 and c_3, such that

$$g(w) - g(w-s) \le \frac{1}{a} g(w) + b(a)\big(g(s)+1\big), \tag{4.23}$$

where one can choose

$$b(a) = \big((2c_2) + \cdots + (2c_2)^{m(a)}\big) + 1/a,$$

and $m(a) = [\log_2(ac_3+1)] + 1$.

PROOF OF LEMMA 4.2. The inequality (4.22) can be derived immediately from (2.5). Meanwhile, (4.23) remains to be shown. For $a > 1$, consider two cases.

Case 1. If $s < w \le (ac_3+1)s$, denote $m := m(a) = [\log_2(ac_3+1)] + 1$. As g is nondecreasing and by (2.5), we have

$$g(w) \le g\big(2^m s\big) \le 2c_2 g\big(2^{m-1}s\big) + c_2.$$

By induction, we have

$$\begin{aligned} g(w) &\le (2c_2)^m g(s) + c_2\big(1 + (2c_2) + \cdots + (2c_2)^{m-1}\big) \\ &\le b(a)\big(g(s)+1\big), \end{aligned} \tag{4.24}$$

where $b(a) = 2c_2(1 + (2c_2) + \cdots + (2c_2)^{m(a)-1}) + 1/a$.

Case 2. If $w > (ac_3+1)s$, by (2.6), we have

$$\begin{aligned} g(w) - g(w-s) &= \int_0^s g'(w-t)\,dt \\ &\le c_3 \int_0^s \frac{1+g(w-t)}{1+(w-t)}\,dt \\ &\le \frac{1}{a}\big(g(w)+1\big). \end{aligned} \tag{4.25}$$

By (4.24) and (4.25), this completes the proof. □

LEMMA 4.3. *For $w \ge 0$ and any $a > 0$, we have*

$$g'(w) \le \frac{1}{a} g(w) + c_3\big(g(ac_3)+1\big) + 1/a. \tag{4.26}$$

PROOF OF LEMMA 4.3. Recall that (2.6) states that for $w \ge 0$,

$$g'(w) \le c_3\left(\frac{1+g(w)}{1+w}\right).$$

Fix $a > 0$. When $w > ac_3$, we have

$$g'(w) \le \frac{1}{a}\big(g(w)+1\big).$$

When $w \le ac_3$, by the monotonicity property of g, we have

$$g'(w) \le c_3\big(g(ac_3)+1\big).$$

This completes the proof. □

For $s > 0$, define

$$f(w,s) = \begin{cases} e^{G(w)-G(w-s)} - 1 & w > s, \\ e^{G(w)} - 1 & 0 \le w \le s, \\ 0 & w \le 0. \end{cases} \tag{4.27}$$

We next consider a ratio property of $f(w,s)$. It is easy to see that $f(w,s)$ is absolutely continuous with respect to both w and s, and the partial derivatives are

$$\begin{aligned} \frac{\partial}{\partial w} f(w,s) &= e^{G(w)-G(w-s)}\big(g(w) - g(w-s)\big) I(w > s) \\ &\quad + e^{G(w)} g(w) I(0 \le w \le s) \end{aligned} \tag{4.28}$$

and

$$\frac{\partial}{\partial s} f(w,s) = e^{G(w)-G(w-s)} g(w-s) I(0 < s \le w). \tag{4.29}$$

LEMMA 4.4. *Let $f(w) := f(w,s)$ be defined as in (4.27). For $0 \le \delta \le 1$ and $\delta|g(w)| \le d_1$, we have*

$$\sup_{|u|\le\delta} \left| \frac{f(w+u)+1}{f(w)+1} \right| I(w+u \ge 0) \le \mu_2, \tag{4.30}$$

where $\mu_2 = \exp(c_2(d_1 + \mu_1) + \mu_1)$. Moreover, we have

$$\sup_{|u|\le\delta} \left| f''(w+u) \right| \le \mu_3\big(g^2(w) + 1\big)\big(f(w) + 1\big), \tag{4.31}$$

where $\mu_3 = 2c_2^2(c_3 + 1)(\mu_1^2 + 1)\mu_2$.

PROOF. Recall that $\mu_1 = \max(g(1), |g(-1)|) + 1$. When $w + u \ge 0$ and $w \ge 0$, as g is nondecreasing, we have

$$\begin{aligned} \sup_{|u|\le\delta} \left| \frac{f(w+u)+1}{f(w)+1} \right| &\le e^{G(w+\delta)-G(w)} \\ &\le e^{\delta|g(w+\delta)|} \le e^{c_2(d_1+\mu_1)}, \end{aligned}$$

where in the last inequality we use (4.22). When $w + u \ge 0$, $w < 0$ and $|u| \le \delta$, we have $0 \le w + u < \delta \le 1$; hence, by the nondecreasing property of g,

$$\sup_{|u|\le\delta} \left| \frac{f(w+u)+1}{f(w)+1} \right| \le \sup_{|u|\le\delta} e^{G(w+u)} \le e^{G(\delta)} \le e^{\mu_1}.$$

This proves (4.30).

For $f''(w)$, by (4.28),

$$\begin{aligned} f''(w) = {} & e^{G(w)-G(w-s)}\big(g(w) - g(w-s)\big)^2 I(w > s) \\ & + e^{G(w)-G(w-s)}\big(g'(w) - g'(w-s)\big) I(w > s) \\ & + e^{G(w)} g^2(w) I(0 \le w \le s) \\ & + e^{G(w)} g'(w) I(0 \le w \le s). \end{aligned}$$

As g is nondecreasing, we have $g'(w-s) \ge 0$; thus, $g'(w) - g'(w-s) \le g'(w)$. For $w > s$, $0 \le g(w) - g(w-s) \le g(w)$. Therefore,

$$f''(w) \le \big(g'(w) + g^2(w)\big)\big(f(w) + 1\big) I(w \ge 0).$$

By (2.6), for $c_3 > 1$, we have

$$g^2(w) + g'(w) \le g^2(w) + c_3\bigl(1 + g(w)\bigr) \le 2(c_3 + 1)\bigl(g^2(w) + 1\bigr).$$

Hence,

$$f''(w) \le 2(c_3 + 1)\bigl(g^2(w) + 1\bigr)\bigl(f(w) + 1\bigr).$$

By (4.22) and (4.30), we have

$$\sup_{|u|\le\delta} \bigl|f''(w+u)\bigr| \le \mu_3\bigl(g^2(w) + 1\bigr)\bigl(f(w) + 1\bigr),$$

where $\mu_3 = 2c_2^2(c_3 + 1)(\mu_1^2 + 1)\mu_2$. This completes the proof of Lemma 4.4. $\square$

Let W be the random variable defined as in Theorem 2.1. For $0 \le \tau \le \max(2, \tau_1 + 1, \tau_2)$ and $s > 0$, Lemmas 4.5 and 4.6 give the properties of $\mathrm{E}|g(W)|^\tau$, $\mathrm{E}|g(W)|^\tau e^{G(W)} I(0 \le W \le s)$ and $\mathrm{E}|g(W)|^\tau e^{G(W)-G(W-s)} I(W > s)$, which play a key role in the proofs of Propositions 4.1 and 4.2.

LEMMA 4.5. *Suppose that conditions (A1)–(A4) and (2.9)–(2.11) are satisfied with $\delta \le 1$. For $0 \le \tau \le \max(2, \tau_1 + 1, \tau_2)$, we have*

$$\mathrm{E}\bigl|g(W)\bigr|^\tau \le C. \tag{4.32}$$

Moreover, for $s > 0$, we have

$$\mathrm{E}\bigl(e^{G(W)-G(W-s)} g^\tau(W) I(W > s)\bigr) \le C\bigl(1 + g^\tau(s)\bigr)\bigl(\mathrm{E}\bigl(f(W,s)\bigr) + 1\bigr), \tag{4.33}$$

and

$$\mathrm{E}\bigl(e^{G(W)} g^\tau(W) I(0 \le W \le s)\bigr) \le C\bigl(1 + g^\tau(s)\bigr)\bigl(\mathrm{E}\bigl(f(W,s)\bigr) + 1\bigr). \tag{4.34}$$

PROOF OF LEMMA 4.5. In this proof, we always assume that $\delta \le 1$.

We first prove (4.32). Without loss of generality, we consider only the case where $\tau \ge 2$. As $\delta|g(W)| \le d_1$, we have $\mathrm{E}|g(W)|^\tau < \infty$. To bound $\mathrm{E}|g(W)|^\tau$, without loss of generality, we consider only $\mathrm{E}g^\tau(W)I(W \ge 0)$. Let $g_+(w) := g(w)I(w \ge 0)$. As $g(0) = 0$ and g is differentiable, we find that $g_+(w)$ is absolutely continuous. By (2.1), we have

$$\begin{aligned}\mathrm{E}\bigl\{g^\tau(W)I(W \ge 0)\bigr\} &= \mathrm{E}\bigl\{g(W) \cdot g_+^{\tau-1}(W)\bigr\} \\ &:= Q_1 + Q_2,\end{aligned} \tag{4.35}$$

where

$$Q_1 = (\tau - 1)\mathrm{E}\int_{|u|\le\delta} g_+^{\tau-2}(W+u) g'(W+u) I(W+u \ge 0) \hat{K}(u)\,du,$$

$$Q_2 = \mathrm{E}R(W) g_+^{\tau-1}(W).$$

The following inequality is well known: for any $a > 0$, $x, y \ge 0$ and $\tau > 1$

$$x^{\tau-1} y \le \frac{\tau-1}{a\tau} x^\tau + \frac{a^{\tau-1}}{\tau} y^\tau. \tag{4.36}$$

For the first term Q_1, by (2.6), we have

$$g'(w+u) \le c_3\bigl(1 + \bigl|g(w+u)\bigr|\bigr).$$

Thus, for $w+u\geq 0$,

$$
\begin{aligned}
&\sup_{|u|\leq\delta} g_+^{\tau-2}(w+u)g'(w+u)\\
&\quad\leq c_3\sup_{|u|\leq\delta}\bigl(g_+^{\tau-1}(w+u)+g_+^{\tau-2}(w+u)\bigr)\\
&\quad\leq 2c_3\sup_{|u|\leq\delta}\bigl(g_+^{\tau-1}(w+u)+1\bigr)\\
&\quad\leq \frac{1-\alpha}{8\times(2c_2)^\tau d_0(\tau-1)}\sup_{|u|\leq\delta}\bigl|g(w+u)\bigr|^\tau+D_{1,0},
\end{aligned}
$$

where we use (4.36) with

$$
a=\frac{8\times(2c_2)^{\tau+1}d_0(\tau-1)}{1-\alpha}\quad\text{and}\quad x=\bigl|g_+(w+u)\bigr|
$$

in the last inequality. Here and in the sequel, $D_{1,0}$, $D_{2,0}$, etcetera denote constants depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. By (4.22), we have

$$
\sup_{|u|\leq\delta}\bigl|g(w+u)\bigr|^\tau\leq(2c_2)^\tau\bigl(\bigl|g(w)\bigr|^\tau+\mu_1^\tau\bigr).
$$

Then, by (2.9), we have

$$
Q_1\leq\frac{1-\alpha}{8}\mathrm{E}\bigl|g(W)\bigr|^\tau+D_{2,0}. \tag{4.37}
$$

For Q_2, by (2.11) and using (4.36) again, we have

$$
Q_2\leq\alpha\mathrm{E}g_+^\tau(W)+\frac{1-\alpha}{4}\mathrm{E}g_+^\tau(W)+\left(\frac{4}{1-\alpha}\right)^{\tau-1}. \tag{4.38}
$$

Hence, by (4.35), (4.37) and (4.38), we have

$$
\mathrm{E}g_+^\tau(W)\leq\frac{1}{6}\mathrm{E}\bigl|g(W)\bigr|^\tau+D_{3,0}.
$$

Similarly, we have

$$
\mathrm{E}g_-^\tau(W)\leq\frac{1}{6}\mathrm{E}\bigl|g(W)\bigr|^\tau+D_{4,0}.
$$

Combining the two foregoing inequalities yields (4.32).

As to (4.33) and (4.34), we first consider the case where $\tau\geq 2$. Write $f(w):=f(w,s)$. By (2.1) and (4.28), we have

$$
\begin{aligned}
\mathrm{E}\bigl(g(W)^\tau f(W)\bigr)&=\mathrm{E}g(W)\bigl\{g(W)^{\tau-1}f(W)\bigr\}\\
&=M_1+M_2+M_3+M_4,
\end{aligned} \tag{4.39}
$$

where

$$
\begin{aligned}
M_1&=\mathrm{E}\int_{|u|\leq\delta} g^\tau(W+u)e^{G(W+u)}I(0\leq W+u\leq s)\hat{K}(u)\,du,\\
M_2&=\mathrm{E}\int_{|u|\leq\delta} g^{\tau-1}(W+u)\bigl(g(W+u)-g(W+u-s)\bigr)\\
&\qquad\times e^{G(W+u)-G(W+u-s)}I(W+u>s)\hat{K}(u)\,du,\\
M_3&=(\tau-1)\mathrm{E}\int_{|u|\leq\delta} g^{\tau-2}(W+u)g'(W+u)f(W+u)\hat{K}(u)\,du,\\
M_4&=\mathrm{E}R(W)g^{\tau-1}(W)f(W).
\end{aligned} \tag{4.40}
$$

We next give the bounds of M_1, M_2, M_3 and M_4. For M_1, by (2.9) and (4.30) and noting that g is nondecreasing, we have

$$\begin{aligned} M_1 &\le d_0 g^\tau(s) \mathrm{E} \sup_{|u|\le\delta} \big(f(W+u)+1\big) I(0 \le W+u \le s) \\ &\le d_0\mu_2 g^\tau(s)\mathrm{E}\big(f(W)+1\big). \end{aligned} \tag{4.41}$$

To bound M_2, we first give the bound of $g(w+u)$ and $g(w+u)-g(w+u-s)$ for $|u|\le\delta$. By (4.22), we have

$$\sup_{|u|\le\delta} \big|g(w+u)\big| \le c_2\big(\big|g(w)\big|+\mu_1\big). \tag{4.42}$$

Furthermore, by (4.23) with $a = 2^{\tau+2} d_0\mu_2 c_2^\tau/(1-\alpha)$, for $w+u>s$, there exists a constant D_1 depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ such that

$$\begin{aligned} &\sup_{|u|\le\delta} \big|g(w+u)-g(w+u-s)\big| \\ &\quad\le \frac{1-\alpha}{2^{\tau+3} d_0\mu_2 c_2^\tau} \sup_{|u|\le\delta}\big|g(w+u)\big| + D_1\big(g(s)+1\big). \end{aligned} \tag{4.43}$$

By (4.36), (4.42) and (4.43), we have

$$\begin{aligned} &\sup_{|u|\le\delta} \big|g(W+u)^{\tau-1}\big(g(W+u)-g(W+u-s)\big)\big| \\ &\quad\le \left(\frac{1-\alpha}{2^{\tau+3} d_0\mu_2 c_2^\tau} \sup_{|u|\le\delta}\big|g(W+u)\big| + D_1\big(g(s)+1\big)\right) \sup_{|u|\le\delta}\big|g(W+u)\big|^{\tau-1} \\ &\quad\le \frac{1-\alpha}{2^{\tau+2} d_0\mu_2 c_2^\tau} \sup_{|u|\le\delta}\big|g(W+u)\big|^\tau + \frac{2^{\tau+3} d_0\mu_2 c_2^\tau}{\tau(1-\alpha)} \times D_1^\tau\big(1+g(s)\big)^\tau \\ &\quad\le \frac{1-\alpha}{4 d_0\mu_2}\big(\big|g(W)\big|^\tau + \mu_1^\tau\big) + \frac{2^{\tau+3} d_0\mu_2 c_2^\tau}{\tau(1-\alpha)} \times D_1^\tau\big(1+g(s)\big)^\tau \\ &\quad\le \frac{1-\alpha}{4 d_0\mu_2}\big|g(W)\big|^\tau + D_2\big(1+g^\tau(s)\big), \end{aligned}$$

where

$$D_2 = \frac{2^{2\tau+3} d_0\mu_2 c_2^\tau}{\tau(1-\alpha)} \times D_1^\tau + \frac{(1-\alpha)\mu_1^\tau}{4 d_0\mu_2}.$$

By (2.9) and (4.30), we have

$$\begin{aligned} M_2 \le{}& \frac{1-\alpha}{4}\mathrm{E}\big|g(W)\big|^\tau\big(f(W)+1\big) \\ &+ d_0\mu_2 D_2\big(1+g^\tau(s)\big)\mathrm{E}\big(f(W)+1\big). \end{aligned} \tag{4.44}$$

For M_3, by Lemma 4.3 and similar to (4.44), we have

$$\begin{aligned} M_3 \le{}& \frac{1-\alpha}{4}\mathrm{E}\big|g(W)\big|^\tau\big(f(W)+1\big) \\ &+ D_3\big(1+g^\tau(s)\big)\mathrm{E}\big(f(W)+1\big), \end{aligned} \tag{4.45}$$

where D_3 is a finite constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ.

For M_4, by (2.11) and (4.36), we have

$$\begin{aligned} M_4 &\le \alpha \mathrm{E}|g(W)|^\tau f(W) + \alpha \mathrm{E}|g(W)|^{\tau-1} f(W) \\ &\le \left(\alpha + \frac{1-\alpha}{4}\right)\mathrm{E}|g(W)|^\tau f(W) + \left(\frac{4\alpha}{1-\alpha}\right)^{\tau-1} \mathrm{E} f(W). \end{aligned} \tag{4.46}$$

By (4.39), (4.41) and (4.44)–(4.46), we have

$$\begin{aligned} \mathrm{E}|g(W)|^\tau f(W) \le{}& \left(\alpha + \frac{3(1-\alpha)}{4}\right)\mathrm{E}|g(W)|^\tau f(W) \\ &+ \left(D_4 + \mathrm{E}|g(W)|^\tau\right)(1 + g^\tau(s))\mathrm{E}\left(f(W)+1\right), \end{aligned}$$

where D_4 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. Rearranging the inequality gives

$$\mathrm{E}|g(W)|^\tau f(W) \le \frac{4(D_4 + \mathrm{E}|g(W)|^\tau)}{1-\alpha}(1 + g^\tau(s))\mathrm{E}\left(f(W)+1\right). \tag{4.47}$$

Combining (4.47) and (4.32), we have

$$\mathrm{E}|g(W)|^\tau \left(f(W)+1\right) \le D_5\left(1 + g^\tau(s)\right)\mathrm{E}\left(f(W)+1\right), \tag{4.48}$$

where D_5 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. This proves (4.33) and (4.34) for $\tau \ge 2$.

For $0 \le \tau < 2$ with $\mathrm{E}|g(W)|^2 < \infty$. By the Cauchy inequality, we have

$$\left(1 + g^{2-\tau}(s)\right)|g(w)|^\tau \le 1 + g^2(s) + 2g^2(w),$$

and noting that for $s > 0$ and $g(s) > 0$,

$$\begin{aligned} |g(w)|^\tau &\le \frac{1 + g^2(s) + 2g^2(w)}{1 + g^{2-\tau}(s)} \\ &\le g^\tau(s) + \frac{1 + 2g^2(w)}{1 + g^{2-\tau}(s)}. \end{aligned} \tag{4.49}$$

By (4.48) with $\tau = 2$, we have

$$\mathrm{E}|g(W)|^2\left(f(W)+1\right) \le D_6\left(1 + g^2(s)\right)\mathrm{E}\left(f(W)+1\right), \tag{4.50}$$

where D_6 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ.

Thus, for $0 \le \tau < 2$, by (4.50) and (4.49), we have

$$\begin{aligned} \mathrm{E}|g(W)|^\tau\left(f(W)+1\right) &\le g^\tau(s)\mathrm{E}\left(f(W)+1\right) \\ &\quad + \frac{\mathrm{E}(f(W)+1) + 2\mathrm{E}g^2(W)(f(W)+1)}{1 + g^{2-\tau}(s)} \\ &\le D_7\left(1 + g^\tau(s)\right)\mathrm{E}\left(f(W)+1\right), \end{aligned}$$

where D_7 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \alpha$ and τ. This completes the proof together with (4.48). □

LEMMA 4.6. *Let $f(w,s)$ be defined as in (4.27). Let $0 < \delta \le 1$ and $s > 0$. Suppose that the conditions in Theorem 2.1 are satisfied. Then, we have*

$$\begin{aligned} &\mathrm{E}\left(f(W,s)+1\right) \\ &\quad \le C(1+s)\exp\{C\left(\delta\left(1 + sg^2(s)\right) + \delta_1\left(1 + sg^{\tau_1+1}(s)\right) \\ &\qquad + \delta_2\left(1 + sg^{\tau_2}(s)\right)\right)\}. \end{aligned} \tag{4.51}$$

REMARK 4.1. Following the proof of Lemma 4.6, if we assume that the condition (2.7) is replaced by (2.18) and (2.19), then the result of Lemma 4.6 still holds.

PROOF OF LEMMA 4.6. Let $h(s) = \mathrm{E} f(W, s)$ and let $f(w) := f(w, s)$. By (4.28) and (4.29), for $s > 0$, we have

$$\begin{aligned} h'(s) &= \mathrm{E}\big(e^{G(W)-G(W-s)} g(W-s) I(W > s)\big) \\ &= \mathrm{E}\big(f(W)g(W)\big) + \mathrm{E}\big(g(W) I(W > 0)\big) - \mathrm{E}\big(f'(W)\big). \end{aligned}$$

We first show that $h'(s)$ can be bounded by a function of $h(s)$. We then solve the differential inequality to obtain the bound of $h(s)$, using an idea similar to that in the proof of Lemma 4.5.

By (2.1), we have

$$\mathrm{E}\big(f(W)g(W)\big) - \mathrm{E}\big(f'(W)\big) = T_1 + T_2 + T_3, \tag{4.52}$$

where

$$\begin{aligned} T_1 &= \mathrm{E}\bigg(\int_{|u|\le\delta} \big(f'(W+u) - f'(W)\big)\hat{K}(u)\,du\bigg), \\ T_2 &= \mathrm{E} f'(W)\big(\mathrm{E}(\hat{K}_1 \mid W) - 1\big), \\ T_3 &= \mathrm{E}\big(f(W)R(W)\big). \end{aligned}$$

We next give the bounds of T_1, T_2 and T_3.

(i) The bound of T_1. By (4.31), we have

$$\begin{aligned} &\sup_{|u|\le\delta} \big|f'(w+u) - f'(w)\big| \\ &\quad\le \delta \sup_{|u|\le\delta} \big|f''(w+u)\big| \\ &\quad\le \delta\mu_3\big(g^2(w) + 1\big)\big(f(w) + 1\big). \end{aligned}$$

By (2.9) and Lemma 4.5, we have

$$\begin{aligned} |T_1| &\le \delta d_0 \mu_3 \mathrm{E}\big(g^2(W) + 1\big)\big(f(W) + 1\big) \\ &\le D_8 \delta\big(1 + g^2(s)\big)\mathrm{E}\big(f(W) + 1\big), \end{aligned} \tag{4.53}$$

where D_8 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1$ and α.

(ii) The bound of T_2. By (2.7) and Lemma 4.5, we have

$$\begin{aligned} |T_2| &\le \delta_1 \mathrm{E}\big(\big|g(W)\big|\big(\big|g(W)\big|^{\tau_1} + 1\big)\big(f(W) + 1\big)\big) \\ &\le 2\delta_1 \mathrm{E}\big(\big|g(W)\big|^{\tau_1+1} + 1\big)\big(f(W) + 1\big) \\ &\le D_9 \delta_1\big(1 + g^{\tau_1+1}(s)\big)\mathrm{E}\big(f(W) + 1\big), \end{aligned} \tag{4.54}$$

where D_9 is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_1$ and α.

(iii) The bound of T_3. By (2.8) and Lemma 4.5, we have

$$\begin{aligned} T_3 &\le \delta_2 \mathrm{E}\big(\big|g(W)\big|^{\tau_2} + 1\big) f(W) \\ &\le D_{10} \delta_2\big(1 + g^{\tau_2}(s)\big)\mathrm{E}\big(f(W) + 1\big), \end{aligned} \tag{4.55}$$

where D_{10} is a constant depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_2$ and α.

By (4.32), we have

$$\mathrm{E}g(W)I(W>0) \le D_{11}, \tag{4.56}$$

where D_{11} is a constant depending on $c_2, c_3, d_0, d_1, \mu_1$ and α. By (4.52)–(4.56), we have

$$\begin{aligned} h'(s) \le D_{11} + D_{12}\big(\delta\big(1+g^2(s)\big) + \delta_1\big(1+g^{\tau_1+1}(s)\big) + \delta_2\big(1+g^{\tau_2}(s)\big)\big) \\ \times \mathrm{E}\big(f(W)+1\big), \end{aligned}$$

where $D_{12} = \max(D_8, D_9, D_{10})$. Therefore,

$$\begin{aligned} h'(s) &\le D_{12}\big(\delta\big(1+g^2(s)\big) + \delta_1\big(1+g^{\tau_1+1}(s)\big) + \delta_2\big(1+g^{\tau_2}(s)\big)\big)h(s) \\ &\quad + D_{11} + D_{12}\big(\delta\big(1+g^2(s)\big) + \delta_1\big(1+g^{\tau_1+1}(s)\big) + \delta_2\big(1+g^{\tau_2}(s)\big)\big). \end{aligned}$$

By solving the differential inequality and given that $s + sg^{\tau}(s) \le 1 + (1+g^{-\tau}(1))sg^{\tau}(s)$ for $\tau > 0$ and $s \ge 0$, we have

$$\begin{aligned} \mathrm{E}\big(f(W)+1\big) \le C_1(1+s)\exp\big\{C_2\big(\delta\big(1+sg^2(s)\big) + \delta_1\big(1+sg^{\tau_1+1}(s)\big) \\ + \delta_2\big(1+sg^{\tau_2}(s)\big)\big)\big\}, \end{aligned}$$

where C_1 and C_2 are constants depending on $c_2, c_3, d_0, d_1, \mu_1, \tau_1, \tau_2$ and α. This completes the proof. □

The next lemma gives the properties of the Stein solution.

LEMMA 4.7. *Let f_z be the solution to Stein's equation (4.3). Then, for $z \ge 0$,*

$$|f_z(w)g(w)| \le \begin{cases} 1-F(z) & w \le 0, \\ F(z) & w > 0, \end{cases} \tag{4.57}$$

$$0 \le f_z(w) \le \begin{cases} (1-F(z))/c_1 & w \le 0, \\ F(z)/c_1 & w > 0, \end{cases} \tag{4.58}$$

and

$$|f_z'(w)| \le \begin{cases} 2\big(1-F(z)\big) & w \le 0, \\ 1 & 0 < w \le z, \\ 2F(z) & w > z. \end{cases} \tag{4.59}$$

PROOF OF LEMMA 4.7. Our first step is to prove (4.57). By (4.4), we have

$$f_z(w)g(w) = \begin{cases} \dfrac{F(w)g(w)(1-F(z))}{p(w)} & w \le z, \\ \dfrac{F(z)g(w)(1-F(w))}{p(w)} & w > z. \end{cases} \tag{4.60}$$

Without loss of generality, we must consider only three cases when $z > 0$:

1. $w < 0$: By (4.2),

$$|f_z(w)g(w)| \le 1 - F(z).$$

2. $0 \le w \le z$: Since $w \le z$, $1 - F(z) \le 1 - F(w)$, thus by (4.1),

$$|f_z(w)g(w)| \le \frac{F(w)|g(w)|(1-F(w))}{p(w)} \le F(w) \le F(z).$$

3. $w > z$: By (4.1),

$$\left|f_z(w)g(w)\right| \le F(z).$$

We can have a similar argument when $z \le 0$, which completes the proof of (4.57). Additionally, (4.58) can be shown similarly. (4.59) follows directly from (4.3) and (4.57). □

LEMMA 4.8. *For $z > 0$ and $0 \le \tau \le \max(2, \tau_1 + 1, \tau_2)$,*

$$\mathrm{E}\big(f_z(W)\left|g(W)\right|^\tau\big) \le C\big(1 + zg^\tau(z)\big)\big(1 - F(z)\big), \tag{4.61}$$

provided that $\max(\delta, \delta_1, \delta_2) \le 1$ *and* $\delta zg^2(z) + \delta_1 zg^{\tau_1+1}(z) + \delta_2 zg^{\tau_2}(z) \le 1$.

PROOF OF LEMMA 4.8. By (4.4),

$$\mathrm{E}\big(f_z(W)\left|g(W)\right|^\tau\big) = T_4 + T_5 + T_6,$$

where

$$T_4 = F(z)\mathrm{E}\bigg(\frac{1 - F(W)}{p(W)}\left|g(W)\right|^\tau I(W > z)\bigg),$$
$$T_5 = \big(1 - F(z)\big)\mathrm{E}\bigg(\frac{F(W)}{p(W)}\left|g(W)\right|^\tau I(W < 0)\bigg),$$
$$T_6 = \big(1 - F(z)\big)\mathrm{E}\bigg(\frac{F(W)}{p(W)}\left|g(W)\right|^\tau I(0 \le W \le z)\bigg).$$

(i) For T_4, we first consider the case when $\tau \ge 1$. As $g(w)$ is increasing, $e^{G(w)-G(w-z)}$ is also increasing with respect to w; thus,

$$I(W > z) \le \frac{e^{G(W)-G(W-z)}I(W > z)}{e^{G(z)}}.$$

By Lemma 4.6, we have $\max(\delta, \delta_1, \delta_2) \le 1$ and z, satisfying that $\delta zg^2(z) + \delta_1 zg^{\tau_1+1}(z) + \delta_2 zg^{\tau_2}(z) \le 1$,

$$\mathrm{E}\big(f(W, z) + 1\big) \le C(1 + z).$$

Hence, by (4.1) and Lemma 4.5, we have

$$\begin{aligned} T_4 &\le Ce^{-G(z)}\mathrm{E}\left|g(W)\right|^{\tau-1}e^{G(W)-G(W-z)}I(W > z) \\ &\le Ce^{-G(z)}\big(1 + g^{\tau-1}(z)\big)\mathrm{E}\big(f(W, z) + 1\big) \\ &\le Ce^{-G(z)}\big(1 + zg^{\tau-1}(z)\big) \\ &\le C\big(1 + zg^\tau(z)\big)\big(1 - F(z)\big), \end{aligned} \tag{4.62}$$

for $\max(\delta, \delta_1, \delta_2) \le 1$ and z, satisfying that $\delta zg^2(z) + \delta_1 zg^{\tau_1+1}(z) + \delta_2 zg^{\tau_2}(z) \le 1$. If $0 \le \tau < 1$, then $g^\tau(w) \le 2(1 + g(w))/(1 + g^{1-\tau}(z))$ for $w > z$. Therefore, (4.62) also holds for $0 \le \tau < 1$.

(ii) As to T_5, because $F(w)/p(w) \le 1/c_1$ for $w \le 0$,

$$T_5 \le \frac{1}{c_1}\big(1 - F(z)\big)\mathrm{E}\left|g(W)\right|^\tau I(W < 0).$$

By (4.32), we have

$$T_5 \le C\big(1 - F(z)\big) \tag{4.63}$$

for some constant C.

(iii) We now bound T_6. By Lemmas 4.6 and 4.5,

$$\begin{aligned}
T_6 &\le C(1-F(z))\mathrm{E}e^{G(W)}|g(W)|^\tau I(0\le W\le z)\\
&\le C(1-F(z))(1+g^\tau(z))\mathrm{E}e^{G(W)}I(0\le W\le z)\\
&\le C(1-F(z))(1+zg^\tau(z)).
\end{aligned}\tag{4.64}$$

By (4.62)–(4.64), we have

$$\mathrm{E}(f_z(W)|g(W)|^\tau)\le C(1+zg^\tau(z))(1-F(z)),$$

which completes the proof. □

4.4. *Proofs of Propositions* 4.1 *and* 4.2. We are now ready to give the proofs of Propositions 4.1 and 4.2.

PROOF OF PROPOSITIONS 4.1. Recalling (2.9), we have

$$\begin{aligned}
I_1 &\le d_0\mathrm{E}\Big(\sup_{|t|\le\delta}|f_z(W+t)g(W+t)-f_z(W)g(W)|\Big)\\
&\le \delta d_0\mathrm{E}\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|.
\end{aligned}\tag{4.65}$$

We first prove (4.6). By Lemma 4.7, $\|f_z\|\le 1/c_1$ and $\|f_z'\|\le 2$. Thus, for $0<\delta\le 1$,

$$\begin{aligned}
&\mathrm{E}\Big(\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|\Big)\\
&\quad\le \mathrm{E}\Big(\sup_{|t|\le\delta}(|f_z(W+t)g'(W+t)|+|f_z'(W+t)g(W+t)|)\Big)\\
&\quad\le (2+1/c_1)\mathrm{E}\Big(\sup_{|t|\le\delta}(|g'(W+t)|+|g(W+t)|)\Big)\\
&\quad\le 4c_3(1+1/c_1)(1+c_2)(\mathrm{E}|g(W)|+\mu_1),
\end{aligned}\tag{4.66}$$

where in the last inequality we use (2.6) and Lemma 4.2. This proves (4.6) by (4.66), (4.65) and (4.32).

Next, we prove (4.7). Similar to the proof of (4.6), we first calculate the following term:

$$\mathrm{E}\Big(\sup_{|t|\le\delta}|(f_z(W+t)g(W+t))'|\Big).$$

Note that

$$\begin{aligned}
&(f_z(w)g(w))'\\
&\quad=\begin{cases}\dfrac{p(w)g(w)+F(w)g'(w)+F(w)g^2(w)}{p(w)}(1-F(z)) & w\le z,\\[2ex] \dfrac{-p(w)g(w)+(1-F(w))g'(w)+(1-F(w))g^2(w)}{p(w)}F(z) & w>z.\end{cases}
\end{aligned}\tag{4.67}$$

For $w+t\le 0$, by (4.2), we have

$$\begin{aligned}
&|(f_z(w+t)g(w+t))'|\\
&\quad\le (1-F(z))\Big(2|g(w+t)|+\frac{g'(w+t)}{\max\{c_1,|g(w+t)|\})}\Big)\\
&\quad\le (1-F(z))(2|g(w+t)|+c_3(1+1/c_1))\\
&\quad\le C(1-F(z))(|g(w)|+1).
\end{aligned}$$

Thus, by (4.32),

$$\mathrm{E}\Big(\sup_{|t|\le\delta}\big|\big(f_z(W+t)g(W+t)\big)'\big|I(W+t\le 0)\Big)\le C\big(1-F(z)\big). \tag{4.68}$$

For $w+t>z$, and $|t|\le\delta$, again by Lemma 4.2, we have

$$\begin{aligned}&\big|\big(f_z(w+t)g(w+t)\big)'\big|\\&\quad\le F(z)\bigg(\big|g(w+t)\big|+\frac{1-F(w+t)}{p(w+t)}\big(\big|g'(w+t)\big|+\big|g(w+t)\big|^2\big)\bigg)\\&\quad\le C\big(1+\big|g(w+t)\big|\big)\\&\quad\le C\big(\big|g(w)\big|+1\big).\end{aligned} \tag{4.69}$$

Hence, by Lemmas 4.5 and 4.6, we have

$$\begin{aligned}&\mathrm{E}\sup_{|t|\le\delta}\big|\big(f_z(W+t)g(W+t)\big)'\big|I(W+t\ge z)\\&\quad\le C\mathrm{E}\big(\big(\big|g(W)\big|+1\big)I(W>z-\delta)\big)\\&\quad\le Cp(z-\delta)\mathrm{E}\big(e^{G(W)-G(W-z+\delta)}\big|g(W)\big|I(W>z-\delta)\big)\\&\quad\le Ce^{\delta g(z)}p(z)\big(1+g(z)\big)\mathrm{E}\big(e^{G(W)-G(W-z+\delta)}I(W>z-\delta)\big)\\&\quad\le Ce^{\delta g(z)}\big(1+zg^2(z)\big)\big(1-F(z)\big),\end{aligned} \tag{4.70}$$

where we use the Lemma 4.1 in the last line. Also note that by (4.17), $\delta g(z)\le\delta+\delta zg^2(z)\le 2$ for $z\ge 1$ and $\delta g(z)\le\mu_1$ for $0\le z\le 1$. Hence,

$$\delta g(z)\le\max(2,\mu_1). \tag{4.71}$$

Thus, (4.70) and (4.71) yield

$$\begin{aligned}&\mathrm{E}\Big(\sup_{|t|\le\delta}\big|\big(f_z(W+t)g(W+t)\big)'\big|I(W+t>z)\Big)\\&\quad\le C\big(1+zg^2(z)\big)\big(1-F(z)\big).\end{aligned} \tag{4.72}$$

For $w+t\in(0,z)$ and $|t|\le\delta$, by (4.22), (4.67) and (4.71), we have

$$\begin{aligned}&\big|\big(f_z(w+t)g(w+t)\big)'\big|\\&\quad\le C\big(1-F(z)\big)e^{G(w+t)}\big(1+g(w+t)^2\big)\\&\quad\le C\big(1-F(z)\big)e^{G(w)+\delta g(z)}\big(1+\big|g(w)\big|^2\big)\\&\quad\le C\big(1-F(z)\big)e^{G(w)}\big(1+\big|g(w)\big|^2\big).\end{aligned} \tag{4.73}$$

By Lemmas 4.5 and 4.6 and (4.22), we have

$$
\begin{aligned}
&\mathrm{E}\Big(\sup_{|t|\le\delta}\big|\big(f_z(W+t)g(W+t)\big)'\big|I(0\le W+t\le z)\Big)\\
&\quad\le C\big(1-F(z)\big)\mathrm{E}e^{G(W)}\big(1+|g(W)|^2\big)I(-\delta\le W\le z+\delta)\\
&\quad= C\big(1-F(z)\big)\mathrm{E}e^{G(W)}\big(1+|g(W)|^2\big)I(-\delta\le W\le 0)\\
&\qquad+C\big(1-F(z)\big)\mathrm{E}e^{G(W)}\big(1+|g(W)|^2\big)I(0\le W\le z+\delta)\\
&\quad\le Ce^{\mu_1}(1+\mu_1^2)\big(1-F(z)\big)\\
&\qquad+C\big(1-F(z)\big)\big(1+(z+\delta)g^2(z+\delta)\big)\\
&\quad\le C\big(1-F(z)\big)\big(1+zg^2(z)\big).
\end{aligned}
\tag{4.74}
$$

Putting together (4.68), (4.72) and (4.74) gives

$$
\mathrm{E}\Big(\sup_{|t|\le\delta}\big|\big(f_z(W+t)g(W+t)\big)'\big|\Big)\le C\big(1+zg^2(z)\big)\big(1-F(z)\big).
\tag{4.75}
$$

By (4.65) and (4.75), we obtain (4.7). □

PROOF OF PROPOSITION 4.2. By Lemma 4.7, we have $\|f_z g\|\le 1$; thus, by (2.7) and (4.32),

$$
I_2+I_3\le C\mathrm{E}\big|\mathrm{E}(\hat{K}_1\mid W)-1\big|\le C\delta_1\big(\mathrm{E}\big(|g(W)|^{\tau_1}\big)+1\big)\le C\delta_1.
$$

To bound I_4, by (2.8), (4.32) and (4.58), we have

$$
I_4\le C\delta_2.
$$

This proves (4.8).

We now move to prove (4.9) and (4.10). As to I_2, by (2.7) and Lemma 4.8, for $z\ge 0$, $\max(\delta,\delta_1,\delta_2)\le 1$ and $\delta zg^2(z)+\delta_1 zg^{\tau_1+1}(z)+\delta_2 zg^{\tau_2}(z)\le 1$, we have

$$
\begin{aligned}
I_2&\le\delta_1\mathrm{E}\big(f_z(W)|g(W)|\big(|g(W)|^{\tau_1}+1\big)\big)\\
&\le C\delta_1\mathrm{E}\big(f_z(W)\big(1+|g(W)|^{\tau_1+1}\big)\big)\\
&\le C\delta_1\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big).
\end{aligned}
\tag{4.76}
$$

As to I_3, note that

$$
I(W>z)\le\frac{e^{G(W)-G(W-z)}}{e^{G(z)}}I(W>z).
$$

By Lemmas 4.5 and 4.6,

$$
\begin{aligned}
&\mathrm{E}\big(\big(1+|g(W)|^{\tau_1}\big)I(W>z)\big)\\
&\quad\le Cp(z)\mathrm{E}\big(e^{G(W)-G(W-z)}\big(1+|g(W)|^{\tau_1}\big)I(W>z)\big)\\
&\quad\le C\big(1+zg^{\tau_1}(z)\big)p(z)\\
&\quad\le C\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big),
\end{aligned}
\tag{4.77}
$$

where we use (4.1) in the last inequality. Thus, by Lemma 4.5 and (4.77),

$$
\begin{aligned}
I_3 &\le \delta_1\big(1-F(z)\big)E\big(\big|g(W)\big|^{\tau_1}+1\big)\\
&\quad+\delta_1\mathrm{E}\big(\big(\big|g(W)\big|^{\tau_1}+1\big)I(W>z+\delta)\big)\\
&\le \delta_1\big(1-F(z)\big)E\big(\big|g(W)\big|^{\tau_1}+1\big)\\
&\quad+\delta_1\mathrm{E}\big(\big(\big|g(W)\big|^{\tau_1}+1\big)I(W>z)\big)\\
&\le C\delta_1\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big).
\end{aligned}
\tag{4.78}
$$

(4.9) now follows by (4.76) and (4.78).

As to I_4, because $|R(W)|\le\delta_2(1+|g(W)|^{\tau_2})$, by (4.61), we have

$$
I_4\le C\delta_2\big(1+zg^{\tau_2}(z)\big)\big(1-F(z)\big).
\tag{4.79}
$$

This completes the proof of Proposition 4.2. □

4.5. *Proof of Remark* 2.1. In this subsection, we assume that the condition (2.7) in Theorem 2.1 is replaced by (2.17)–(2.19), then the result of Remark 2.1 follows from the proof of Theorem 2.1, Propositions 4.1 and 4.2 and the following proposition:

PROPOSITION 4.3. *Assume that the condition* (2.7) *in Theorem* 2.1 *is replaced by* (2.17)–(2.19), *then* (4.8) *and* (4.9) *hold*.

PROOF OF PROPOSITIONS 4.3. Following the proof of Propositions 4.2, it suffices to prove the following inequalities:

$$
\mathrm{E}|K_2|\le\delta_1,
\tag{4.80}
$$

and for $z>0$ such that $\delta zg^2(z)+\delta_1 zg^{\tau_1+1}(z)+\delta_2 zg^{\tau_2}(z)\le 1$,

$$
\mathrm{E}\big|f_z(W)g(W)K_2\big|\le C\delta_1\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big),
\tag{4.81}
$$

$$
\mathrm{E}|K_2|I(W>z)\le C\delta_1\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big).
\tag{4.82}
$$

For (4.80), by (2.19) with $s=0$, noting that $\zeta(W,0)\equiv 1$ and $g(0)=0$, we have (4.80) holds.

For (4.81), by the definition of f_z, and by Lemmas 4.1 and 4.7, we have

$$
\mathrm{E}\big|f_z(W)g(W)K_2\big|\le T_7+T_8+T_9,
\tag{4.83}
$$

where

$$
\begin{aligned}
T_7&=\big(1-F(z)\big)\mathrm{E}|K_2|I(W<0),\\
T_8&=\big(1-F(z)\big)\mathrm{E}|K_2|g(W)e^{G(W)}I(0\le W\le z),\\
T_9&=\mathrm{E}|K_2|I(W>z).
\end{aligned}
$$

For T_7, by (4.80), we have

$$
T_7\le\delta_1\big(1-F(z)\big).
\tag{4.84}
$$

For T_8, by the monotonicity of $g(\cdot)$ and by (2.19) and Remark 4.1, we have

$$
\begin{aligned}
T_8&\le\big(1-F(z)\big)g(z)\mathrm{E}|K_2|\zeta(W,z)\\
&\le\delta_1\big(1-F(z)\big)\big(1+g(z)^{\tau_1+1}\big)\mathrm{E}\zeta(W,s)\\
&\le C\delta_1\big(1+zg^{\tau_1+1}(z)\big)\big(1-F(z)\big).
\end{aligned}
\tag{4.85}
$$

For T_9, by the Chebyshev inequality, by (2.19) and Lemmmas 4.6 and 4.1, we have

$$
\begin{aligned}
T_9 &\le e^{-G(z)}\mathrm{E}|K_2|\zeta(W,z)I(W>z)\\
&\le C\delta_1\bigl(1+zg^{\tau_1}(z)\bigr)e^{-G(z)} \qquad (4.86)\\
&\le C\delta_1\bigl(1+zg^{\tau_1+1}(z)\bigr)\bigl(1-F(z)\bigr).
\end{aligned}
$$

The inequality (4.81) follows from (4.83)–(4.86) while (4.82) follows from (4.86). This completes the proof. □

4.6. *Proof of Remark* 2.2. In this subsection, we assume that the condition (2.11) is replaced by (2.20) and (2.21). The conclusion of Remark 2.2 follows from the proof of Theorem 2.1 and the following lemma.

LEMMA 4.9. *Let the conditions in Remark* 2.2 *be satisfied. Furthermore*, $0<\delta\le 1$, *and* $s>0$ *such that* $\delta s g^2(s)\le 1$. *For* $0\le\tau\le\max\{2,\tau_1+1,\tau_2\}$, *inequalities* (4.32)–(4.34) *hold.*

PROOF. Recall that $s_0=\max\{s:\delta s g^2(s)\le 1\}$ and $\delta\le 1$. We have

$$
s_0\ge s_1 \quad\text{and}\quad \delta s_1 g^2(s_1)=1.
$$

Following the proof of Lemma 4.5, it suffices to prove the following two inequalities.

For Q_2 defined in (4.35),

$$
Q_2\le\left(\alpha+\frac{1-\alpha}{4}\right)\mathrm{E}g_+^\tau(W)+C, \tag{4.87}
$$

and for M_4 defined in (4.40),

$$
M_4\le\left(\alpha+\frac{1-\alpha}{4}\right)\mathrm{E}\bigl|g(W)\bigr|^\tau f(W)+\left(\frac{4\alpha}{1-\alpha}\right)^{\tau-1}\mathrm{E}f(W)+C. \tag{4.88}
$$

For Q_2, by (2.20) and similar to (4.38), we have

$$
Q_2\le\left(\alpha+\frac{1-\alpha}{4}\right)\mathrm{E}g_+^\tau(W)+\left(\frac{4}{1-\alpha}\right)^{\tau-1}+d_2\mathrm{E}g_+^\tau(W)I(W>\kappa).
$$

For the last term, by (2.10) and (2.21) and noting that $0\le\tau\le\max\{2,\tau_1+1,\tau_2\}$, we obtain

$$
\begin{aligned}
d_2\mathrm{E}g_+^\tau(W)I(W>\kappa)&\le d_1^{-\tau}d_2\delta^{-\tau}\mathrm{P}(W>\kappa)\\
&\le d_1^{-\tau}d_3\delta^{-\tau}\exp\bigl(-2s_0d_1^{-1}\delta^{-1}\bigr)\\
&\le d_1^{-\tau}d_3\sup_{\delta>0}\bigl\{\delta^{-\tau}\exp\bigl(-2s_1d_1^{-1}\delta^{-1}\bigr)\bigr\} \qquad (4.89)\\
&=d_3\left(\frac{\tau}{2s_1}\right)^\tau e^{-\tau},
\end{aligned}
$$

where the equality holds when $\delta=2s_1/(d_1\tau)$. The inequality (4.87) follows from (4.38) and (4.89).

As to M_4, by (2.20), we have

$$
\begin{aligned}
M_4\le{}&\left(\alpha+\frac{1-\alpha}{4}\right)\mathrm{E}\bigl|g(W)\bigr|^\tau f(W)+\left(\frac{4\alpha}{1-\alpha}\right)^{\tau-1}\mathrm{E}f(W)\\
&+d_2\mathrm{E}\bigl|g^\tau(W)\bigr|e^{G(W)-G(W-s)}I(W>\kappa).
\end{aligned}
$$

For the last term, by (2.10) and (2.21) and noting that $g(\cdot)$ is nondecreasing and $s \le s_0$, similar to (4.89), we have

$$\begin{aligned} d_2 \mathrm{E}\big|g^\tau(W)\big|e^{G(W)-G(W-s)} I(W>\kappa) &\le d_1^{-\tau} d_2 \delta^{-\tau} e^{s d_1^{-1}\delta^{-1}} \mathrm{P}(W>\kappa) \\ &\le d_1^{-\tau} d_3 \delta^{-\tau} e^{-s_0 d_1^{-1}\delta^{-1}} \\ &\le d_1^{-\tau} d_3 \sup_{\delta>0}\{\delta^{-\tau} e^{-s_1 d_1^{-1}\delta^{-1}}\} \\ &= d_3\left(\frac{\tau}{s_1}\right)^\tau e^{-\tau}, \end{aligned} \tag{4.90}$$

where the equality holds when $\delta = s_1/(d_1\tau)$. Combining (4.46) and (4.90), inequality (4.88) holds. Following the proof of Lemma 4.5 and replacing (4.38) and (4.46) with (4.87) and (4.88), respectively, we complete the proof of Lemma 4.9. □

5. Proofs of Theorems 3.1–3.2.

5.1. *Proof of Theorem* 3.1. In this subsection, we use Remarks 2.1 and 2.2 to prove the result.

We first prove some preliminary lemmas.

LEMMA 5.1. *Let $\xi \sim \rho$. For $s \in \mathbb{R}$, define*

$$\psi_n(s) = \frac{\mathrm{E}(\xi e^{\frac{\xi^2}{2n}+\xi s})}{\mathrm{E}(e^{\frac{\xi^2}{2n}+\xi s})}, \qquad \psi_\infty(s) = \frac{\mathrm{E}(\xi e^{\xi s})}{\mathrm{E}(e^{\xi s})},$$

and

$$\phi_n(s) = \frac{\mathrm{E}(\xi^2 e^{\frac{\xi^2}{2n}+\xi s})}{\mathrm{E}(e^{\frac{\xi^2}{2n}+\xi s})}, \qquad \phi_\infty(s) = \frac{\mathrm{E}(\xi^2 e^{\xi s})}{\mathrm{E}(e^{\xi s})}.$$

Let $m = \frac{1}{n}\sum_{i=1}^n X_i$ and $m_i = \frac{1}{n}\sum_{j\ne i} X_j$. We have for each $1 \le i \le n$,

$$\big|\psi_\infty(m) - \psi_n(m_i)\big| \le Cn^{-1}, \tag{5.1}$$

$$\big|\phi_\infty(m) - \phi_n(m_i)\big| \le Cn^{-1}, \tag{5.2}$$

where C is a positive constant depending only on L.

PROOF OF LEMMA 5.1. Recall that $|\xi| \le L$ and observe that

$$\begin{gathered} \big|\mathrm{E}(\xi(e^{\frac{\xi^2}{2n}+\xi s} - e^{\xi s}))\big| \le \frac{1}{2n}\mathrm{E}|\xi|^3 e^{\frac{\xi^2}{2n}+\xi s} \le \frac{L^3}{2n} e^{L^2/2}\mathrm{E}e^{\xi s}, \\ \big|\mathrm{E}(e^{\frac{\xi^2}{2n}+\xi s} - e^{\xi s})\big| \le \frac{1}{2n}\mathrm{E}|\xi|^2 e^{\frac{\xi^2}{2n}+\xi s} \le \frac{L^2}{2n} e^{L^2/2}\mathrm{E}e^{\xi s}, \\ \big|\mathrm{E}\xi e^{\xi s}\big| \le L\mathrm{E}e^{\xi s}, \end{gathered}$$

and

$$\mathrm{E}(e^{\frac{\xi^2}{2n}+\xi s}) \ge \mathrm{E}e^{\xi s}.$$

Hence,

$$
\begin{aligned}
\left|\psi_n(s)-\psi_\infty(s)\right| &\le \frac{|\mathrm{E}e^{\xi s}|\times|\mathrm{E}\xi e^{\frac{\xi^2}{2n}+\xi s}-\mathrm{E}\xi e^{\xi s}|}{\mathrm{E}e^{\frac{\xi^2}{2n}+\xi s}\mathrm{E}e^{\xi s}} \\
&\quad + \frac{|\mathrm{E}\xi e^{\xi s}|\times|\mathrm{E}e^{\frac{\xi^2}{2n}+\xi s}-\mathrm{E}e^{\xi s}|}{\mathrm{E}e^{\frac{\xi^2}{2n}+\xi s}\mathrm{E}e^{\xi s}} \\
&\le Cn^{-1},
\end{aligned}
\tag{5.3}
$$

where $C>0$ depends only on L. Moreover,

$$\psi_\infty'(s)=\frac{\mathrm{E}(\xi^2 e^{\xi s})}{\mathrm{E}(e^{\xi s})}-\left\{\frac{\mathrm{E}(\xi e^{\xi s})}{\mathrm{E}(e^{\xi s})}\right\}^2.$$

Recalling that $|\xi|\le L$, $|X_i|\le L$ and $|m-m_i|\le L/n$, and using the fact that

$$\sup_{|s|\le L}\left|\psi_\infty'(s)\right|\le L^2,$$

we have

$$\left|\psi_\infty(m)-\psi_\infty(m_i)\right|\le L^3 n^{-1}. \tag{5.4}$$

Following (5.3)–(5.4), the inequality (5.1) holds.

A similar argument implies that (5.2) holds as well. $\square$

Set

$$\mathcal{F}=\sigma\{X_1,\dots,X_n\}. \tag{5.5}$$

For any $1\le i,j\le n$, define

$$\mathcal{F}^{(i)}=\sigma\bigl(\{X_k,k\ne i\}\bigr),\mathcal{F}^{(i,j)}=\sigma\bigl(\{X_k,k\ne i,j\}\bigr). \tag{5.6}$$

LEMMA 5.2. *Let* $W=n^{-1+\frac{1}{2k}}\sum_{i=1}^n X_i$, $G(w)=h^{(2k)}(0)w^{2k}/(2k)!$, *and*

$$\zeta(w,s)=\begin{cases} e^{G(w)-G(w-s)} & w>s,\\ e^{G(w)} & 0\le w\le s,\\ 1 & w<0.\end{cases}$$

Suppose (2.9), (2.10), (2.20) *and* (2.21) *are satisfied. Then, we have*

$$\mathrm{E}\left|\frac{1}{n}\sum_{i=1}^n\bigl(X_i^2-\mathrm{E}\bigl(X_i^2\mid\mathcal{F}^{(i)}\bigr)\bigr)\right|\zeta(W,s)\le Cn^{-1/k}\bigl(1+|s|^2\bigr)\mathrm{E}\zeta(W,s), \tag{5.7}$$

where C is a positive constant depending only on ρ.

We are now ready to prove Theorem 3.1.

PROOF OF THEOREM 3.1. We first construct the exchangeable pair of W. For each $1\le i\le n$, let X_i' follow the conditional distribution of X_i given $\{X_j,j\ne i\}$, and be conditionally independent of X_i given $\{X_j,j\ne i\}$. Let I be a random index uniformly distributed among $\{1,2,\dots,n\}$, independent of all other random variables. Define $S_n'=S_n-X_I+X_I'$ and

$W' = n^{-\frac{1}{2k}} S_n'$. Then (W, W') is an exchangeable pair. Let $\mathcal{F}$, $\mathcal{F}^{(i)}$ and $\mathcal{F}^{(i,j)}$ be defined as in (5.5) and (5.6). Let $\psi_n, \psi_\infty, \phi_n$ and ϕ_∞ be as defined in Lemma 5.1. We have

$$\mathrm{E}\big(X_i' \mid \mathcal{F}^{(i)}\big) = \mathrm{E}\big(X_i \mid \mathcal{F}^{(i)}\big) = \psi_n\big(m_i(\mathbf{X})\big), \tag{5.8}$$

where $m_i(\mathbf{X}) = \frac{1}{n}\sum_{j\neq i} X_j$.

Thus,

$$\begin{aligned}\mathrm{E}\big(X_I - X_I' \mid \mathcal{F}\big) &= \frac{1}{n}\sum_{i=1}^n \mathrm{E}\big(X_i - X_i' \mid \mathcal{F}\big)\\ &= m(\mathbf{X}) - \frac{1}{n}\sum_{i=1}^n \mathrm{E}\big(X_i' \mid \mathcal{F}^{(i)}\big)\\ &= m(\mathbf{X}) - \frac{1}{n}\sum_{i=1}^n \psi_n\big(m_i(\mathbf{X})\big)\\ &= m(\mathbf{X}) - \psi_\infty\big(m(\mathbf{X})\big) + r(\mathbf{X})\\ &= h'\big(m(\mathbf{X})\big) + r(\mathbf{X}),\end{aligned} \tag{5.9}$$

where $m(\mathbf{X}) = (1/n)\sum_{i=1}^n X_i$, h is as defined in (3.3), and

$$r(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^n \big\{\psi_\infty\big(m(\mathbf{X})\big) - \psi_n\big(m_i(\mathbf{X})\big)\big\}. \tag{5.10}$$

By Lemma 5.1, we have

$$\big|r(\mathbf{X})\big| \le Cn^{-1},$$

where $C > 0$ is a constant depending only on ρ. As ρ is symmetric, $h^{(2k+1)}(0) = 0$. By the Taylor expansion, for $|w| \le L$,

$$\big|h'(w) - g(w)\big| \le C|w|^{2k+1},$$

where $C > 0$ is a constant depending only on L. Therefore,

$$\begin{aligned}\mathrm{E}\big(W - W' \mid \mathcal{F}\big) &= n^{-1+\frac{1}{2k}} \mathrm{E}\big(X_I - X_I' \mid \mathcal{F}\big)\\ &= n^{-1+\frac{1}{2k}}\big(h'\big(m(\mathbf{X})\big) + r\big(m(\mathbf{X})\big)\big)\\ &= \lambda\big(g(W) + R(W)\big),\end{aligned}$$

where $\lambda = n^{-2+1/k}$,

$$g(w) = \frac{h^{(2k)}(0)}{(2k-1)!} w^{2k-1}, \qquad \big|R(w)\big| \le C_1 n^{-1/k}\big(|w|^{2k+1} + 1\big),$$

where $C_1 > 0$ depends only on ρ.

We now check the conditions (2.20) and (2.21). As $g(w) = \frac{h^{(2k)}(0)}{(k-1)!} w^{2k-1}$, then

$$\big|R(W)\big| \le C_1\bigg(\frac{(k-1)!}{h^{(2k)}(0)} + 1\bigg) n^{-1/k}\big(\big|W^2 g(W)\big| + 1\big).$$

Moreover, recalling that $|W| \le Ln^{\frac{1}{2k}}$, we have

$$\big|R(W)\big| \le C_1\big(n^{(2k-1)/2k} L^{2k+1} + 1\big).$$

Set

$$\kappa = \Big(2C_1\Big(1+\frac{(k-1)!}{h^{(2k)}(0)}\Big)\Big)^{-1/2} n^{\frac{1}{2k}},$$

where $d_2 = C_1(n^{(2k-1)/2k}L^{2k+1} + \frac{(k-1)!}{h^{(2k)}(0)} + 2)$. Thus,

$$|R(W)| \le \frac{1}{2}\big(|g(W)|+1\big) + d_2 I\big(|W|\ge \kappa\big). \tag{5.11}$$

By Chatterjee and Dey [8], Propostion 6, for any $n \ge 1$ and $t \ge 0$,

$$\mathrm{P}\big(|W|\ge t\big) \le 2e^{-c_\rho t^{2k}},$$

where $c_\rho > 0$ is a constant depending only on ρ. Note that $\delta = Ln^{-1+\frac{1}{2k}}$ and by the definition of $g(\cdot)$, we have

$$s_0 = \max\{s : \delta s g^2(s) \le 1\} = C_2 n^{(2k-1)/(2k(4k-1))},$$

where $C_2 > 0$ is a constant depending on ρ. Moreover, there exists a constant $d_1 > 0$ depending on ρ such that $\delta|g(W)| \le d_1$. Then, there exist positive constants C_3 and C_4 depending on ρ such that

$$d_2 e^{2s_0 d_1^{-1}\delta^{-1}}\mathrm{P}\big(|W|\ge\kappa\big) \le C_3(n+1)\exp\{C_4 n^{2(k-1)/(4k-1)} - c_\rho n\} \le d_3, \tag{5.12}$$

where $d_3 > 0$ is a constant depending on ρ. Thus the conditions (2.10), (2.20) and (2.21) hold.

For the conditional second moment, by Lemma 5.1, we have

$$\begin{aligned}
&\mathrm{E}\big((X_I - X_I')^2 \mid \mathcal{F}\big)\\
&\quad= \frac{1}{n}\sum_{i=1}^n \mathrm{E}\big((X_i - X_i')^2 \mid \mathcal{F}\big)\\
&\quad= \frac{1}{n}\sum_{i=1}^n X_i^2 - \frac{2}{n}\sum_{i=1}^n X_i\psi_n\big(m_i(\mathbf{X})\big) + \frac{1}{n}\sum_{i=1}^n \phi_n\big(m_i(\mathbf{X})\big)\\
&\quad= \frac{1}{n}\sum_{i=1}^n \big(X_i^2 - \phi_n\big(m_i(\mathbf{X})\big)\big) - 2m(\mathbf{X})\psi_\infty\big(m(\mathbf{X})\big)\\
&\qquad + 2\phi_\infty\big(m(\mathbf{X})\big) + r_2(\mathbf{X}),
\end{aligned} \tag{5.13}$$

where ψ_n, ϕ_n and ϕ_∞ are as defined in Lemma 5.1. By the Taylor expansion, we have

$$\big|\phi_\infty\big(m(\mathbf{X})\big) - 1\big| = \big|h''\big(m(\mathbf{X})\big)\big| \le Cn^{-1+1/k}\big(1+|W|^{2k-2}\big), \tag{5.14}$$

and

$$\big|m(\mathbf{X})\psi_\infty\big(m(\mathbf{X})\big)\big| \le Cn^{-1/k}|W|^2 + Cn^{-1}|W|^{2k}, \tag{5.15}$$

where $C > 0$ is a constant depending only on ρ. By the definition of (W, W') and (5.13)–(5.15), with $\lambda = n^{-2+1/k}$, we have

$$\begin{aligned}
&\left|\frac{1}{2\lambda}\mathrm{E}\big((W-W')^2 \mid \mathcal{F}\big) - 1\right|\\
&\quad= \left|\frac{1}{2}\mathrm{E}\big((X_I - X_I')^2 \mid \mathcal{F}\big) - 1\right|\\
&\quad\le \frac{1}{2}\left|\frac{1}{n}\sum_{i=1}^n \big(X_i^2 - \phi_n\big(m_i(\mathbf{X})\big)\big)\right| + Cn^{-1/k}\big(1+|W|^2\big).
\end{aligned}$$

Moreover, as $|X_i| \le L$, we have

$$\left|\frac{1}{2\lambda}\mathrm{E}\big((W-W')^2 \mid \mathcal{F}\big)-1\right| \le 2L^2+1 =: d_0.$$

Then (2.9) holds. By Lemma 5.2, we have the condition (2.19) in Remark 2.1 is satisfied.

Hence, we have (2.8)–(2.10) and the conditions in Remarks 2.1 and 2.2 are satisfied with $\delta_1 = \delta_2 = Cn^{-1/k}$, $\tau_1 = \frac{2}{2k-1}$, and $\tau_2 = 1 + \frac{2}{2k-1}$. By Remarks 2.1 and 2.2, we complete the proof of Theorem 3.1. □

It suffices to proof Lemma 5.2.

PROOF OF LEMMA 5.2. In this proof, we denote C by a general positive constant depending only on ρ. By the Cauchy inequality, we have

$$\begin{aligned}
&\mathrm{E}\left|\frac{1}{n}\sum_{i=1}^{n}\big(X_i^2-\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big)\big)\right|\zeta(W,s) \\
&\quad \le \left(\mathrm{E}\left|\frac{1}{n}\sum_{i=1}^{n}\big(X_i^2-\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big)\big)\right|^2 \zeta(W,s) \times \mathrm{E}\zeta(W,s)\right)^{1/2}.
\end{aligned} \tag{5.16}$$

Expand the square term, and we have

$$\mathrm{E}\left|\frac{1}{n}\sum_{i=1}^{n}\big(X_i^2-\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big)\big)\right|^2 \zeta(W,s) = H_1 + H_2, \tag{5.17}$$

where

$$\begin{aligned}
H_1 &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\big\{\big(X_i^2-\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big)\big)^2\zeta(W,s)\big\}, \\
H_2 &= \frac{1}{n^2}\sum_{i\ne j}\mathrm{E}\big\{\big(X_i^2-\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big)\big)\big(X_j^2-\mathrm{E}\big(X_j^2 \mid \mathcal{F}^{(j)}\big)\big)\zeta(W,s)\big\}.
\end{aligned}$$

Recalling that $|X_i| \le L$, we have

$$H_1 \le 4L^4 n^{-1}\mathrm{E}\zeta(W,s). \tag{5.18}$$

As for H_2, we first introduce some notation. For $i \ne j$, let $\mathrm{E}^{(i,j)}$ denote the conditional expectation given $\mathcal{F}^{(i,j)}$, where $\mathcal{F}^{(i,j)}$ is as in (5.6). Note that

$$\mathrm{E}^{(i,j)}\big(X_i^2\big) = \frac{\iint x^2 \exp\big(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij}\big)\,d\rho(x)\,d\rho(y)}{\iint \exp\big(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij}\big)\,d\rho(x)\,d\rho(y)},$$

where $m_{ij} := m_{ij}(\mathbf{X}) = \frac{1}{n}\sum_{k\ne i,j} X_k$. Similar to Lemma 5.1, we have for any $i \ne j$,

$$\big|\mathrm{E}\big(X_i^2 \mid \mathcal{F}^{(i)}\big) - \mathrm{E}^{(i,j)}\big(X_i^2\big)\big| \le Cn^{-1}, \tag{5.19}$$

where $C > 0$ depends only on L. Define

$$H_3 = \frac{1}{n^2}\sum_{i\ne j}\mathrm{E}\big\{\big(X_i^2-\mathrm{E}^{(i,j)}\big(X_i^2\big)\big)\big(X_j^2-\mathrm{E}^{(i,j)}\big(X_j^2\big)\big)\zeta(W,s)\big\}, \tag{5.20}$$

and then by (5.19) and (5.20), we have

$$|H_2 - H_3| \le Cn^{-1}\mathrm{E}\zeta(W,s). \tag{5.21}$$

We now move to give the bound of H_3. Define

$$W^{(i,j)} = W - n^{-1+\frac{1}{2k}}(X_i + X_j).$$

Let

$$q(w,s) = \begin{cases} G(w) - G(w-s) & w > s, \\ G(w) & 0 \le w \le s, \\ 0 & w < 0, \end{cases}$$

and then $q(w,s) = \log \zeta(w,s)$ and $q'(w)$ is continuous on $\mathbb{R}$. Therefore, by the Taylor expansion, we have

$$\begin{aligned} q(W) - q(W^{(i,j)}) &= (W - W^{(i,j)})q'(W^{(i,j)}) \\ &\quad + \frac{1}{2}(W - W^{(i,j)})^2 q''(w_0), \end{aligned} \tag{5.22}$$

where w_0 belongs to either $(W, W^{(i,j)})$ or $(W^{(i,j)}, W)$. Note that $G(w) = Cw^{2k}$ for some constant C, $|W| \le Ln^{\frac{1}{2k}}$ and $|W - W^{(i,j)}| \le 2Ln^{-1+\frac{1}{2k}}$. By the definition of q, we have

$$\begin{aligned} &|(W - W^{(i,j)})q'(W^{(i,j)})| \\ &\quad \le Cn^{-1+\frac{1}{2k}}|W^{(i,j)}|^{2k-1} \\ &\quad \le Cn^{-1+\frac{1}{2k}}(|W|^{2k-1} + 1) \end{aligned} \tag{5.23}$$

and

$$\left|\frac{1}{2}(W - W^{(i,j)})^2 q''(w_0)\right| \le Cn^{-1}, \tag{5.24}$$

where C depends only on ρ. Therefore, by (5.22)–(5.24) and using the fact that $|W| \le Ln^{\frac{1}{2k}}$, we have

$$\begin{aligned} |q(W) - q(W^{(i,j)})| &\le Cn^{-1+\frac{1}{2k}}(|W|^{2k-1} + 1) \\ &\le C. \end{aligned} \tag{5.25}$$

Observe that

$$\mathrm{E}^{(i,j)}\{(X_i^2 - \mathrm{E}^{(i,j)}(X_i^2))(X_j^2 - \mathrm{E}^{(i,j)}(X_j^2))\zeta(W,s)\} = \zeta(W^{(i,j)})M^{(i,j)}, \tag{5.26}$$

where

$$M^{(i,j)} = \mathrm{E}^{(i,j)}\{(X_i^2 - \mathrm{E}^{(i,j)}(X_i^2))(X_j^2 - \mathrm{E}^{(i,j)}(X_j^2))e^{q(W) - q(W^{(i,j)})}\}.$$

Applying the Taylor expansion to the exponential function, we have

$$M^{(i,j)} = M_1^{(i,j)} + M_2^{(i,j)} + M_3^{(i,j)}, \tag{5.27}$$

where

$$\begin{aligned} M_1^{(i,j)} &= \mathrm{E}^{(i,j)}\{(X_i^2 - \mathrm{E}^{(i,j)}X_i^2)(X_j^2 - \mathrm{E}^{(i,j)}X_j^2)\}, \\ M_2^{(i,j)} &= \mathrm{E}^{(i,j)}((X_i^2 - \mathrm{E}^{(i,j)}X_i^2)(X_j^2 - \mathrm{E}^{(i,j)}X_j^2)\{q(W) - q(W^{(i,j)})\}), \end{aligned}$$

and

$$M_3^{(i,j)} = M^{(i,j)} - M_1^{(i,j)} - M_2^{(i,j)}.$$

For $M_1^{(i,j)}$, since $\mathrm{E}^{(i,j)} X_i^2 = \mathrm{E}^{(i,j)} X_j^2$, we have

$$\begin{aligned}
M_1^{(i,j)} &= \mathrm{E}^{(i,j)} X_i^2 X_j^2 - \mathrm{E}^{(i,j)} X_i^2 \mathrm{E}^{(i,j)} X_j^2 \\
&= \frac{\iint x^2 y^2 \exp(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij})\, d\rho(x)\, d\rho(y)}{\iint \exp(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij})\, d\rho(x)\, d\rho(y)} \\
&\quad - \biggl(\frac{\iint x^2 \exp(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij})\, d\rho(x)\, d\rho(y)}{\iint \exp(\frac{1}{2n}(x+y)^2 + (x+y)m_{ij})\, d\rho(x)\, d\rho(y)}\biggr)^2 \\
&= M_{11}^{(i,j)} + M_{12}^{(i,j)},
\end{aligned}$$

where

$$\begin{aligned}
M_{11}^{(i,j)} &= \frac{\iint x^2 y^2 \exp((x+y)m_{ij})\, d\rho(x)\, d\rho(y)}{\iint \exp((x+y)m_{ij})\, d\rho(x)\, d\rho(y)} \\
&\quad - \biggl(\frac{\iint x^2 \exp((x+y)m_{ij})\, d\rho(x)\, d\rho(y)}{\iint \exp((x+y)m_{ij})\, d\rho(x)\, d\rho(y)}\biggr)^2 \\
&= 0,
\end{aligned} \tag{5.28}$$

and $M_{12}^{(i,j)} = M_1^{(i,j)} - M_{11}^{(i,j)}$. Similar to Lemma 5.1, we have

$$|M_{12}^{(i,j)}| \le C n^{-1}. \tag{5.29}$$

By (5.28)–(5.29), we have

$$|M_1^{(i,j)}| \le C n^{-1}. \tag{5.30}$$

For $M_2^{(i,j)}$, by (5.22) and (5.24), we have

$$M_2^{(i,j)} = M_{21}^{(i,j)} + M_{22}^{(i,j)},$$

where

$$\begin{aligned}
M_{21}^{(i,j)} &= n^{-1+\frac{1}{2k}} q'(W^{(i,j)}) \mathrm{E}^{(i,j)} \{(X_i^2 - \mathrm{E}^{(i,j)} X_i^2)(X_j^2 - \mathrm{E}^{(i,j)} X_j^2)(X_i + X_j)\}, \\
M_{22}^{(i,j)} &= \frac{1}{2} \mathrm{E}^{(i,j)} \{(X_i^2 - \mathrm{E}^{(i,j)} X_i^2)(X_j^2 - \mathrm{E}^{(i,j)} X_j^2)(W - W^{(i,j)})^2 q''(w_0)\},
\end{aligned}$$

and w_0 is as defined in (5.22). By (5.24), and recalling that $|X_i| \le L$, we have

$$|M_{22}^{(i,j)}| \le C n^{-1}.$$

Similar to (5.30), we have

$$|\mathrm{E}^{(i,j)} \{(X_i^2 - \mathrm{E}^{(i,j)} X_i^2)(X_j^2 - \mathrm{E}^{(i,j)} X_j^2)(X_i + X_j)\}| \le C n^{-1}.$$

Moreover, recalling that $|W^{(i,j)}| \le L n^{\frac{1}{2k}}$ and $|q'(W^{(i,j)})| \le C n^{1-\frac{1}{2k}}$, we have

$$|M_{21}^{(i,j)}| \le C n^{-1}.$$

Thus,

$$|M_2^{(i,j)}| \le C n^{-1}. \tag{5.31}$$

For $M_3^{(i,j)}$, by the Taylor expansion, noting again that $k \geq 2$, $|W| \leq Ln^{\frac{1}{2k}}$ and $|X_i| \leq L$ for $1 \leq i \leq n$, and by (5.23) and (5.24), we have

$$\begin{aligned}|M_3^{(i,j)}| &\leq C\big|q(W) - q\big(W^{(i,j)}\big)\big|^2 e^{q(W)-q(W^{(i,j)})} \\ &\leq Cn^{-2+1/k}\big(|W|^{4k-2}+1\big) \\ &\leq Cn^{-2/k}\big(|W|^4+1\big).\end{aligned} \tag{5.32}$$

By (5.27) and (5.30)–(5.32), we have

$$\big|M^{(i,j)}\big| \leq Cn^{-2/k}\big(|W|^4+1\big),$$

substituting which to (5.26), we have

$$\begin{aligned}&\mathrm{E}\big|\mathrm{E}^{(i,j)}\big\{\big(X_i^2 - \mathrm{E}^{(i,j)}(X_i^2)\big)\big(X_j^2 - \mathrm{E}^{(i,j)}(X_j^2)\big)\zeta(W,s)\big\}\big| \\ &\quad\leq Cn^{-2/k}\mathrm{E}\big\{\big(|W|^4+1\big)\zeta\big(W^{(i,j)}\big)\big\} \\ &\quad\leq Cn^{-2/k}\mathrm{E}\big\{\big(|W|^4+1\big)\zeta(W,s)\big\} \\ &\quad\leq Cn^{-2/k}(1+s^4)\mathrm{E}\zeta(W,s),\end{aligned} \tag{5.33}$$

where in the last inequality we used Lemma 4.9 recalling the fact that (5.11) and (5.12) are satisfied. By (5.33), we have the term H_3 in (5.20) can be bounded by

$$|H_3| \leq Cn^{-2/k}(1+s^4)\mathrm{E}\zeta(W,s). \tag{5.34}$$

By (5.16)–(5.18), (5.21) and (5.34), we complete the proof of (5.7). $\square$

5.2. *Proof of Theorem* 3.2. In this subsection, we use Remark 2.2 to prove the result.

PROOF OF THEOREM 3.2. For any $\sigma \in \Sigma$, $uv \in D$ and $s, t \in \{0, 1\}$, let σ_{uv}^{st} denote the configuration $\tau \in \Sigma$, such that $\tau_i = \sigma_i$ for $i \neq u, v$ and $\tau_u = s$, $\tau_v = t$. Let (σ_u', σ_v') be independent of (σ_u, σ_v) and follow the conditional distribution

$$\mathrm{P}\big(\sigma_u' = s, \sigma_v' = t \mid \sigma\big) = \frac{p(\sigma_{uv}^{st})}{\sum_{s,t\in\{0,1\}} p(\sigma_{uv}^{st})}.$$

Let $M = \sum_{i=1}^n \sigma_i$ and $M' = M - \sigma_u - \sigma_v + \sigma_u' + \sigma_v'$. Then, by Chen [15], (M, M') is exchangeable. Also, by Chen [15], Proposition 2, we have

$$\mathrm{E}\big(M - M' \mid \sigma\big) = L_1\big(m(\sigma)\big) + R_1\big(m(\sigma)\big), \tag{5.35}$$

$$\mathrm{E}\big((M - M')^2 \mid \sigma\big) = L_2\big(m(\sigma)\big) + R_2\big(m(\sigma)\big), \tag{5.36}$$

where $m(\sigma) = M/n$ and

$$L_1(x) = \frac{2(1-x)(x^2 - (1-x)e^{2\tau(x)})}{(1-x) + e^{2\tau(x)}}, \quad \text{for } 0 < x < 1,$$

$$L_2(x) = \frac{4(1-x)(x^2 + (1-x)e^{2\tau(x)})}{(1-x) + e^{2\tau(x)}}, \quad \text{for } 0 < x < 1,$$

$$\big|R_1(x)\big| + \big|R_2(x)\big| \leq \frac{C}{n}$$

for some constant C. Next, we consider two cases. In the first case, $(J, h) \notin \Gamma \cup \{(J_c, h_c)\}$, and in the second case, $(J, h) = (J_c, h_c)$.

Case 1. When $(J,h)\notin\Gamma\cup\{(J_c,h_c)\}$. Define $W=n^{-1/2}(M-nm_0)$ and $W'=n^{-1/2}(M'-nm_0)$; then, (W,W') is also an exchangeable pair. Moreover,

$$|W-W'|\le 2n^{-1/2}=:\delta.$$

Note that $L_1(m_0)=0$ by observing $m_0^2=(1-m_0)e^{2\tau(m_0)}$. Moreover, we have

$$L_1'(m_0)=\frac{1}{2\lambda_0}L_2(m_0)>0,$$

where $\lambda_0=(-1/H''(m_0))-(1/2J)>0$. By the Taylor expansion, we have

$$L_1\bigl(m(\sigma)\bigr)=L_1'(m_0)\bigl(m(\sigma)-m_0\bigr)+\int_{m_0}^{m(\sigma)}L_1''(s)\bigl(m(\sigma)-s\bigr)\,ds.$$

Let $\lambda=L_2(m_0)/(2n)$, and we have

$$n^{-1/2}L_1\bigl(m(\sigma)\bigr)=\lambda\bigl(\lambda_0^{-1}W+r(W)\bigr),$$

where

$$r(W)=2n^{1/2}L_2^{-1}(m_0)\int_{m_0}^{m(\sigma)}L_1''(s)\bigl(m(\sigma)-s\bigr)\,ds.$$

Therefore, together with the definition of (W,W') and (5.35), we have

$$\mathrm{E}\bigl(W-W'\mid W\bigr)=n^{-1/2}\bigl(L_1\bigl(m(\sigma)\bigr)+R_1\bigl(m(\sigma)\bigr)\bigr)=\lambda\bigl(g(W)+R(W)\bigr),$$

where

$$g(W)=W/\lambda_0\quad\text{and}\quad R(W)=r(W)+\frac{2n^{1/2}}{L_2(m_0)}R_1\bigl(m(\sigma)\bigr).$$

Thus, conditions (A1)–(A4) hold for $g(w)=w/\lambda_0$. Furthermore, $\delta|g(W)|\le 2/\lambda_0$, as $n^{-1/2}|W|\le 1$.

By Chen [15], Lemma 1, there exist constants $C_0,C_1>0$ such that

$$|R(W)|\le C_0n^{-1/2}\bigl(W^2+1\bigr),\tag{5.37}$$

and

$$\left|\frac{1}{2\lambda}\mathrm{E}\bigl((W-W')^2\mid W\bigr)-1\right|\le C_1n^{-1/2}\bigl(|W|+1\bigr)$$

and $|\hat{K}_1|=\frac{\Delta^2}{2\lambda}\le 4/L_2(m_0)$. Therefore, (2.7)–(2.10) are satisfied with $\tau_1=1,\tau_2=2,\delta_1=\delta_2=O(1)n^{-1/2}$ and $d_0=4/L_2(m_0)$ and $d_1=2/\lambda_0$.

It suffices to prove (2.20)–(2.21). By (5.37), we have for $|W|\le\frac{\sqrt{n}}{2\lambda_0C_0}$,

$$|R(W)|\le\frac{1}{2}\bigl(|g(W)|+1\bigr),\tag{5.38}$$

and for $|W|>\frac{\sqrt{n}}{2\lambda_0C_0}$, recalling that $|W|\le 1$, we have $|R(W)|\le C_0(\sqrt{n}+1)$. Then, (2.20) holds with $\alpha=1/2, d_2=C_0(\sqrt{n}+1)$ and $\kappa=\sqrt{n}/(2\lambda_0C_0)$. By Chen [15], Lemma 2, when $(J,h)\notin\Gamma\cup\{(J_c,h_c)\}$, for any $u>0$, there exists a constant $\eta>0$ such that

$$\mathrm{P}\bigl(|m(\sigma)-m_0|\ge u\bigr)\le Ce^{-n\eta}$$

for some constant C. Hence,

$$d_2\mathrm{P}\bigl(|W|>\kappa\bigr)\le C(\sqrt{n}+1)e^{-n\eta}.$$

Note that $s_0 = \max\{s : \delta s g^2(s) \le 1\}$, $g(w) = w/\lambda_0$, $d_1 = \frac{2}{\lambda_0}$ and $\delta = 2n^{-1/2}$, then $s_0 = (\lambda_0/2)^{1/3} n^{1/6}$. Therefore, (2.21) is satisfied. By Remark 2.2, we have

$$\frac{\mathrm{P}(W > z)}{\mathrm{P}(Z_0 > z)} = 1 + O(1)n^{-1/2}(1 + z^3)$$

for $0 \le z \le n^{1/6}$.

Case 2. When $(J, h) = (J_c, h_c)$. Define $W = n^{-3/4}(M - nm_c)$ and $W' = n^{-3/4}(M' - nm_c)$; then, (W, W') is an exchangeable pair. By (5.35), we have

$$\mathrm{E}(W - W'|W) = n^{-3/4}\big(L_1\big(m(\sigma)\big) + R_1\big(m(\sigma)\big)\big).$$

By Chen [15], page 14, we have

$$L_1(m_c) = L_1'(m_c) = L_1''(m_c) = 0, \qquad L_1^{(3)}(m_c) = \frac{\lambda_c}{2} L_2(m_c),$$

where λ_c is given in (3.8). Then, by the Taylor expansion, we have

$$L_1\big(m(\sigma)\big) = \frac{L_1^{(3)}(m_c)}{6}\big(m(\sigma) - m_c\big)^3 + \frac{1}{6}\int_{m_c}^{m(\sigma)} L_1^{(4)}(s)\big(m(\sigma) - s\big)^3\, ds.$$

Then, taking $\lambda = L_2(m_c)/(2n^{3/2})$, by Chen [15], Lemma 1, we have

$$\mathrm{E}\big(W - W' \mid W\big) = \lambda\big(g(W) + R(W)\big),$$

where $g(W) = (\lambda_c/6)W^3$ and

$$R(W) = \frac{n^{3/4}}{2L_2(m_c)}\int_{m_0}^{m(\sigma)} L_1^{(4)}(s)\big(m(\sigma) - m_c\big)^3\, ds + \frac{2n^{3/4}}{L_2(m_c)} R_1(W).$$

Hence, $G(w) = \frac{\lambda_c}{24} w^4$. Based again on Chen [15], Lemma 1, for some constant C, we have

$$|R(W)| \le Cn^{-1/4}\big(|W|^4 + 1\big) \le Cn^{-1/4}\big(|g(W)|^{4/3} + 1\big), \tag{5.39}$$

and

$$\left|\frac{1}{2\lambda}\mathrm{E}\big((W - W')^2 \mid W\big) - 1\right| \le Cn^{-1/4}\big(|g(W)|^{1/3} + 1\big).$$

As $|W - W'| \le 2n^{-3/4}$ and $|W| \le Cn^{1/4}$, it follows that there exist constants d_0 and d_1 such that $n^{-3/4}|g(W)| \le d_1$ and $\hat{K}_1 = (W - W')^2/(2\lambda) \le d_0$. Thus, (2.9) and (2.10) are satisfied. Furthermore, (2.7) and (2.8) hold with $\delta = 2n^{-3/4}$, $\delta_1 = \delta_2 = O(1)n^{-1/4}$ and $\tau_1 = 1/3$, $\tau_2 = 4/3$. It suffices to show that (2.20) and (2.21) are satisfied. By (5.39), there exists a constant $c > 0$ such that for $|W| \le cn^{1/4}$,

$$|R(W)| \le \frac{1}{2}\big(|g(W)| + 1\big).$$

For $|W| \ge cn^{1/4}$, noting that $|W| \le Cn^{1/4}$, we have $|R(W)| \le Cn^{3/4}$. Thus, (2.20) is satisfied with $\alpha = 1/2$, $d_2 = Cn^{3/4}$ and $\kappa = cn^{1/4}$. Furthermore, as $\delta = 2n^{-3/4}$ and $g(w) = (\lambda_c/6)w^3$, we have $s_0 = (18/\lambda_c)^{1/7} n^{3/28}$. In addition, by Chen [15], Lemma 2, when $(J, h) = (J_c, h_c)$, for any $u > 0$, there exists a constant $\eta > 0$ such that

$$\mathrm{P}\big(|m(\sigma) - m_c| \ge u\big) \le Ce^{-n\eta}.$$

Thus,

$$d_2 \mathrm{P}\big(|W| \ge \kappa\big) \le Cn^{3/4} e^{-n\eta} \le Ce^{-2s_0 d_1^{-1}\delta^{-1}}.$$

Then, (2.21) holds. By Remark 2.2, we complete the proof of Theorem 3.2. $\square$

Acknowledgements. We would like to thank the referees for their helpful comments which led to a much improved presentation of the paper.

The first author was partially supported by Hong Kong RGC GRF 14302515 and 14304917. The third author was partially supported by Singapore Ministry of Education Academic Research Fund MOE 2018-T2-076.

REFERENCES

[1] ALBERICI, D., CONTUCCI, P., FEDELE, M. and MINGIONE, E. (2016). Limit theorems for monomer-dimer mean-field models with attractive potential. *Comm. Math. Phys.* **346** 781–799. MR3537336 https://doi.org/10.1007/s00220-015-2543-1

[2] ALBERICI, D., CONTUCCI, P. and MINGIONE, E. (2014). A mean-field monomer-dimer model with attractive interaction: Exact solution and rigorous results. *J. Math. Phys.* **55** 063301. MR3390668 https://doi.org/10.1063/1.4881725

[3] BARBOUR, A. D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659 https://doi.org/10.1007/BF01197887

[4] CAN, V. H. and PHAM, V.-H. (2017). A Cramér type moderate deviation theorem for the critical Curie–Weiss model. *Electron. Commun. Probab.* **22** 62. MR3724560 https://doi.org/10.1214/17-ecp96

[5] CHANG, T. S. (1939). Statistical theory of the adsorption of double molecules. *Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci.* **169** 512–531.

[6] CHATTERJEE, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. MR2435859 https://doi.org/10.1214/07-AOP370

[7] CHATTERJEE, S. (2014). A short survey of Stein's method. In *Proceedings of the International Congress of Mathematicians—Seoul* 2014. *Vol. IV* 1–24. Kyung Moon Sa, Seoul. MR3727600

[8] CHATTERJEE, S. and DEY, P. S. (2010). Applications of Stein's method for concentration inequalities. *Ann. Probab.* **38** 2443–2485. MR2683635 https://doi.org/10.1214/10-AOP542

[9] CHATTERJEE, S. and SHAO, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964 https://doi.org/10.1214/10-AAP712

[10] CHEN, L. H. Y., FANG, X. and SHAO, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293. MR3059199 https://doi.org/10.1214/12-AOP746

[11] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method. Probability and Its Applications* (*New York*). Springer, Heidelberg. MR2732624 https://doi.org/10.1007/978-3-642-15007-4

[12] CHEN, L. H. Y. and SHAO, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028. MR2073183 https://doi.org/10.1214/009117904000000450

[13] CHEN, L. H. Y. and SHAO, Q.-M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** 581–599. MR2331265 https://doi.org/10.3150/07-BEJ5164

[14] CHEN, L. H. Y. and RÖLLIN, A. (2010). Stein couplings for normal approximation. Available at arXiv:1003.6039.

[15] CHEN, W.-K. (2016). Limit theorems in the imitative monomer-dimer mean-field model via Stein's method. *J. Math. Phys.* **57** 083302. MR3535688 https://doi.org/10.1063/1.4960673

[16] DIACONIS, P. (1977). Finite forms of de Finetti's theorem on exchangeability. *Synthese* **36** 271–281. MR0517222 https://doi.org/10.1007/BF00486116

[17] ELLIS, R. S. and NEWMAN, C. M. (1978). The statistics of Curie–Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332 https://doi.org/10.1007/BF01012508

[18] FOWLER, R. H. and RUSHBROOKE, G. S. (1937). An attempt to extend the statistical theory of perfect solutions. *Trans. Faraday Soc.* **33** 1272–1294.

[19] GOLDSTEIN, L. and REINERT, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. MR1484792 https://doi.org/10.1214/aoap/1043862419

[20] LINNIK, J. V. (1961). On the probability of large deviations for the sums of independent variables. In *Proc.* 4*th Berkeley Sympos. Math. Statist. and Prob.*, *Vol. II* 289–306. Univ. California Press, Berkeley, CA. MR0137142

[21] NOURDIN, I. and PECCATI, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. MR2520122 https://doi.org/10.1007/s00440-008-0162-x

[22] NOURDIN, I. and PECCATI, G. (2012). *Normal Approximations with Malliavin Calculus*: *From Stein's method to universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge Univ. Press, Cambridge. MR2962301 https://doi.org/10.1017/CBO9781139084659

[23] PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82. MR0388499

[24] ROBERTS, J. K. (1938). Some properties of mobile and immobile adsorbed films. *Proc*. *Camb*. *Philos*. *Soc*. **34** 399.

[25] SHAO, Q.-M. and ZHANG, Z.-S. (2016). Identifying the limiting distribution by a general approach of Stein's method. *Sci*. *China Math*. **59** 2379–2392. MR3578962 https://doi.org/10.1007/s11425-016-0322-3

[26] SHAO, Q.-M. and ZHANG, Z.-S. (2019). Berry–Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs. *Ann*. *Probab*. **47** 61–108. MR3909966 https://doi.org/10.1214/18-AOP1255

[27] SIMON, B. and GRIFFITHS, R. B. (1973). The $(\phi^4)_2$ field theory as a classical Ising model. *Comm*. *Math*. *Phys*. **33** 145–164. MR0428998

[28] STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ*. *California*, *Berkeley*, *Calif*., 1970/1971), *Vol*. *II*: *Probability Theory* 583–602. MR0402873

[29] STEIN, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **7**. IMS, Hayward, CA. MR0882007